

> 0 <
01/10 Intelligent
> 0 <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "consensus-ags" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.

Selected sequence key from "consens.key":
consensus (AA) ID consensus AA preliminary pattern

1 followed by

```

2 c
2 any character from 9 to 13 times
2 c
2 q or e or r or k
2 any character
2 a or f or g or i or l or m or p or t or v or w or y
2 l or i or v or m
2 any character from 15 to 39 times
2 cc
2 c or d or e or h or k or n or q or r or s
2 q or e or h
2 l or v
2 any character
2 any character
2 any character
2 any character
2 any character
2 c
2 any character
2 any character
2 l or i
2 any character from 13 to 56 times
2 g
2 any character from 15 to 26 times
2 c
2 any character
2 v or i or l or m
2 any character from 1 to 8 times
2 c

```

Selected files:

File : consensus-ags.pep

-- Output Parameters --

Format Options:	File Options:	
Nucleic acid code matching	Exact	No
Find non-matching hits only	Indirect file	No
Report key used	Sequence or key file	No
Note position of hit	List of hits	Yes
Display full annotations	Hit display	Yes
Sequence context	Name and annotations	Yes
	50	

-- Run Parameters --

Run mode	Batch
Time to start comparison	now
Notify at end of run	No

1 match found in sequence:
aap91891 ; Part of the sequence of the Brazil nut 2S-albumin as encoded in the
(from "consensus-ags.pep")
TOIG of: aap91891 check: 9273 from: 1 to: 140

```

ID AAP91891 standard; protein: 140 AA.
XX
XX AAP91891;
XX 29-APR-1990 (first entry)
XX
XX Part of the sequence of the Brazil nut 2S-albumin as encoded in the
DE PBN2S1 plasmid.
XX
XX 2S-albumin: brazil nut; PBN2S1; storage protein gene;
KW heterologous polypeptide.
XX
XX Brazil nut.
OS
XX
XX Key Location/Qualifiers
XX FH Peptide 1..30
XX FT /note="signal peptide" 31..38
XX FT protein
XX FT region 39..43
XX FT /note="mature small subunit"
XX FT /note="processing site" 44..136
XX FT protein /note="mature large subunit"
XX
XX W08903887-A.
XX
XX 05-MAY-1989.
XX
XX 20-OCT-1988; 88W0-EP00944.
XX
XX 20-OCT-1987; 87EP-0402348.
XX
XX (PLAN-) PLANT GENETIC SYST.
XX
XX Vandekerckhove JS, Krebbers E, Botterman J, Leemans J;
XX WP1: 1989-150783/20.
XX DR N-PSDB; AAN91699.
XX
XX Recombinant DNA expression in plants
XX PT - using modified storage protein genes for expressing
XX PT heterologous polypeptide(s) in the seeds
XX
XX Fig 4; ; 121pp; English.
XX
XX The entire 2S-albumin storage protein precursor including
XX CC signal peptide. It is to be inserted into plants under the control of
XX CC a seed-specific promoter and expressed at high levels only or mostly
XX CC in the seed forming stage and produced mostly in the seeds.
XX
XX Sequence 140 AA;
XX

```

AAP91891 Length: 140 September 12, 2002 12:06 Type: P Check: 9273
Found using 'consensus' (consens.key)

1 AAAALVLMALGHATAFRATVTTVEEENQEECREOMROOMLSHCXYRQOMESP
34

61 QTMRRRGMEPHMSECCQLEGMDESCRCRGIRMMMMRQOEMOPRGOMRMMPLAENI

121 PSRCNLSPMRCPMGSGIAGF
131

1 match found in sequence:
aar33390 ; Albumin-type protein encoded by pig8.
(from "consensus-ags.pep")
TOIG of: aar33390 check: 8158 from: 1 to: 175

```

ID  AAR33390 standard; Protein; 175 AA.
XX  AAR33390;
AC
XX  12-JUL-1993 (first entry)
DT
XX  Albumin-type protein encoded by pig8.
DE
XX  Primer: polymerase chain reaction; PCR; cloning; radish; 2S; albumin;
KM  OMB20; pBA3; OMB18; oilseed rape; napin; PNAPI; BamHI; XbaI; amplify;
KM  Raphanus sativus; PEMBL18+; pig8; albumin-type protein; antifungal;
KM  antibacterial; activity; pathogen; plant; thionin.
XX  Raphanus sativus.
OS
XX  Key Location/Qualifiers
FH Peptide 1..21 "Signal peptide"
FT /note="38..72
FT Protein /note="Mature small subunit"
FT 92..173
FT Protein /note="Mature large subunit"
XX
XX  WO9304586-A.
PN
XX  18-MAR-1993.
PD
XX  28-AUG-1992; 92WO-GB01574.
PF
XX  02-SEP-1991; 91GB-0018730.
PR
XX  (ICIL ) IMPERIAL CHEM IND PLC.
PA
XX  Broekaert WF, Cammue BPA, Osborn RW, Rees SB, Terras FRG;
PI  Vanderleyden J;
XX  WPI; 1993-100565/12.
XX  N-P-SDB; AAO38731.
DR
XX  Antimicrobial albumin-type proteins - used for combating fungi
PT  and bacteria on plants and for obtaining DNA for producing
PT  resistant plants
XX
XX  Claim 17; Fig 10; 58pp; English.
PS
XX  This peptide is encoded by the plasmid pig8. pig8 was cloned using
CC  the primer sequences given in AAO38726-27, and represents the radish 2S
CC  albumin cDNA. The sequence of primer OMB20 was based on the published
CC  sequence of the radish 2S albumin pBA3 cDNA. Primer OMB18 was based
CC  on the sequence of the oilseed rape napin PNAPI cDNA. Using these
CC  primers introduces a BamHI recognition site at the 5' end and an XbaI
CC  site at the 3' end. Total cDNA isolated from Raphanus sativus plants
CC  was used as a template for the amplification reaction. The PCR-product
CC  was cut with BamHI and XbaI, subcloned into PEMBL18+ pre-digested with
CC  the same enzymes and sequenced. The resulting plasmid was designated
CC  pig8. This albumin-type protein exhibits antifungal and antibacterial
CC  activity and can be used to eradicate a pathogen already established
CC  on a plant or to protect a plant from future pathogenic attack. Plants
CC  can be transformed using pig8 to provide improved resistance to
CC  pathogens. This albumin-type protein showed synergistic activity with
CC  thionins and can potentiate the antifungal activity of thionins by up
CC  to 70-fold.
XX
SQ  Sequence 175 AA;
AAR33390 Length: 175 September 12, 2002 12:06 Type: P Check: 8158 ..
Found using 'consensus' (consens.key)

```

```

1  MANKLFLVSATLAFFLTLTNASTIRTVVEFDEDDATNPAGPFRIPKRRKFFQQAQHLIRAC
47

```

```

61  QOWLHKQAMQVRGSPSLALDGEFDEDDMENPORPPLQOCNCELHQEPLCAVCPTLKGA
-----
121  SKAVKQIQQQGQQAQHMVMSRIYOTATHLPRVCNIPQVSCPFQKTMGPY
164
-----
1 match found in sequence:
aars3579; Synthetic 2S seed storage protein.
(from "consensus_ags.pep")
TOIG of: aars3579 Check: 9502 from: 1 to: 171
ID  AAR53579 standard; Protein; 171 AA.
XX  AAR53579;
AC
XX  25-OCT-1994 (first entry)
DT
XX  Synthetic 2S seed storage protein.
DE
XX  Nutrition; non-conserved DNA; assembly; crude; degenerate;
KM  oligonucleotides.
XX  Synthetic.
OS
XX  Key Location/Qualifiers
FH Peptide 39..74
FT /note="small subunit"
FT Protein 85..170
FT /note="large subunit"
FT Peptide 1..38
FT /note="signal peptide"
FT 75..84
FT Peptide /note="linker excised as protein folds"
FT Misc-difference 171 /note="excised at protein maturity"
XX
XX  WO9410315-A.
PN
XX  11-MAY-1994.
PD
XX  22-OCT-1993; 93WO-US10090.
PF
XX  23-OCT-1992; 92US-0965664.
PR
XX  (PION-) PIONEER HI-BRED INT INC.
PA
XX  Ballo B;
PI  WPI; 1994-167470/20.
XX  N-P-SDB; AAO63710.
DR
XX  Improving the content of an amino acid in a seed storage protein
PT  to enhance nutritional value -- by replacing non-conserved DNA
PT  with DNA encoding the amino acid
XX
XX  Disclosure; Fig 1; 33pp; English.
PS
XX  The sequence is that of a 2S seed storage protein with increased
CC  lysine content. The protein is synthesised from a set of partial
CC  DNA sequences capable of being assembled in a complementary over-
CC  lapping relationship to provide the complete DNA. The protein is
CC  translated as the straight chain shown and undergoes post translational
CC  modification to give the mature protein which has two subunits joined by
CC  a di-sulphide bond.
XX
SQ  Sequence 171 AA;
AAR53579 Length: 171 September 12, 2002 12:06 Type: P Check: 9502 ..
Found using 'consensus' (consens.key)

```


1 MAKILFFATLALFVLLANASIQTTVEVEDEEDNQLMRCOFPLQORLRACQRTIHR
40

61 AAFSGPDELEDEVEDDENQPRPALRCQNLROVDRCVCLRQAAGQVLRQI

121 ICGPQRLRLFDARNLPLNICIPNICACPRACP
149

1 match found in sequence:
aaW23586; Mabinlin MBLI from Capparis masakal.
(from "consensus_ags.pep")
TOIG of: aaW23586 check: 7526 from: 1 to: 158

ID AAW23586 standard; Protein; 158 AA.
XX
AC AAW23586;
XX
DT 30-SEP-1997 (first entry)
XX
DE Mabinlin MBLI from Capparis masakal.
XX
KM MBL; sweetener; transgenic plant; recombinant protein; food;
KM beverage; animal feed; chewing gum; dental hygiene product;
KM pharmaceutical.
XX
OS Capparis masakal.
XX
FH Key Location/Qualifiers
FT Peptide 1..39
FT /label= Signal
FT Protein 40..158
FT /label= MBLI
FT Region 40..70
FT /label= A_chain
FT Region 71..86
FT /note= "14 amino acids evidently lost when mature
FT dimer is formed"
FT Region 87..158
FT /label= B_chain
XX
PN WO9700945-A1.
XX
PD 09-JAN-1997.
XX
PE 21-JUN-1996; 96WO-US10669.
XX
PR 23-JUN-1995; 95US-0000480.
XX
PA (UYHA-) UNIV HAWAII.
XX
PI Chen H, Hu Z, Sun SSM, Xiong L;
XX
DR WPI: 1997-087372/08.
DR N-PSDB: AAT60775.
XX
PT Mabinlin nucleotide sequences - for prodn. of transgenic plants or
XX recombinant proteins, useful as sweeteners, partic. in foods
PS Disclosure: Fig 3; 38pp; English.
XX
CC The present sequence represents the mabinlin MBLI. The encoding cDNA can
CC modulate the expression of the gene encoding MBL in a cell or tissue.
CC Mabinlin comes from the seeds of the Capparis masakal plant, and
CC is limited by its native availability, and so methods for producing
CC recombinant mabinlin protein are being produced using chimeric genes
CC such as MBLI (see AAT60776). The products provide recombinant materials
CC for the production of the MBL protein in practical amounts, and for the

CC production of transgenic plants containing inherently sweet edible parts
CC by virtue of production of MBL in situ. The MBL protein can be used to
CC sweeten foods, beverages, animal feeds, chewing gum, dental hygiene
CC products and pharmaceuticals.
XX

SO Sequence 158 AA;
AAW23586 Length: 158 September 12, 2002 12:06 Type: P Check: 7526
Found using 'consensus' (consens.key)

1 MAKILITLALFVLLANASITRTVELEDEEDNQPLCRQFOOHRLACQRTYR
42

61 RRAQGLVDELLEDEVEDEEDENQPRPALRLCCNLROVKKPCVCLRQAAGQVLRQI

121 YQGIGSPQVROLERARNLPLNICKIPAVGCOFTRW
153

1 match found in sequence:
aaW23587; Mabinlin MBLI from Capparis masakal.
(from "consensus_ags.pep")
TOIG of: aaW23587 check: 6568 from: 1 to: 155

ID AAW23587 standard; Protein; 155 AA.
XX
AC AAW23587;
XX
DT 30-SEP-1997 (first entry)
XX
DE Mabinlin MBLI from Capparis masakal.
XX
KM MBL; sweetener; transgenic plant; recombinant protein; food;
KM beverage; animal feed; chewing gum; dental hygiene product;
KM pharmaceutical.
XX
OS Capparis masakal.
XX
FH Key Location/Qualifiers
FT Peptide 1..35
FT /label= Signal
FT Protein 36..155
FT /label= MBLI
FT Region 36..68
FT /label= A_chain
FT Region 69..82
FT /note= "14 amino acids evidently lost when mature
FT dimer is formed"
FT Region 83..154
FT /label= B_chain
FT Misc-difference 155
FT /note= "Proline is apparently cleaved off when the
FT mature dimer is formed"
XX
PN WO9700945-A1.
XX
PD 09-JAN-1997.
XX
PE 21-JUN-1996; 96WO-US10669.
XX
PR 23-JUN-1995; 95US-0000480.
XX
PA (UYHA-) UNIV HAWAII.
XX
PI Chen H, Hu Z, Sun SSM, Xiong L;
XX
DR WPI: 1997-087372/08.
DR N-PSDB: AAT60776.
XX

PT Mabinlin nucleotide sequences - for prodn. of transgenic plants or
 PT recombinant proteins, useful as sweeteners, partic. in foods
 XX
 PS Disclosure; Fig 3; 38pp; English.
 XX
 CC The present sequence represents the mabinlin MBLII. The encoding cDNA
 CC can modulate the expression of the gene encoding MBL in a cell or
 CC tissue. Mabinlin comes from the seeds of the Capparis masakakai plant,
 CC and is limited by its native availability, and so methods for producing
 CC recombinant mabinlin protein are being produced using chimeric genes
 CC such as MBLII (see AAT60776). The products provide recombinant materials
 CC for the production of the MBL protein in practical amounts, and for the
 CC production of transgenic plants containing inherently sweet edible parts
 CC by virtue of production of MBL in situ. The MBL protein can be used to
 CC sweeten foods, beverages, animal feeds, chewing gum, dental hygiene
 CC products and pharmaceuticals.
 CC
 SQ Sequence 155 AA;
 AAW23587 Length: 155 September 12, 2002 12:06 Type: P Check: 6568 ..
 Found using 'consensus' (consens.key)
 1 MAKILFLATLALFVLLANASIQTVVEVDEEDNDQMRQORFLQHRLRACQRTHRR
 40
 61 AQRGGGDELEDEVEDNDENQRRPALRQCCNQLRQVDRPCVPLRQAQVLOROI
 121 ICGFOQLRLFLDARNLPNICNIPNIGACPFRAWP
 149
 1 match found in sequence:
 aaw23588; Mabinlin MBLIII from Capparis masakakai.
 (from "consensus-ags.pep")
 TOIG of: aaw23588 check: 8278 from: 1 to: 158
 ID AAW23588 standard; Protein: 158 AA.
 XX
 AC AAW23588;
 XX
 DT 30-SEP-1997 (first entry)
 XX
 DE Mabinlin MBLIII from Capparis masakakai.
 XX
 KM MBL; sweetener; transgenic plant; recombinant protein; food;
 KM beverage; animal feed; chewing gum; dental hygiene product;
 KM pharmaceutical.
 XX
 OS Capparis masakakai.
 XX
 FH Key location/Qualifiers
 FT Peptide 1..39
 FT Protein /label= Signal
 FT Protein /label= MBLIII
 FT Region 40..70
 FT Region /label= A_chain
 FT Region 71..86
 FT /note= "14 amino acids evidently lost when mature
 FT dimer is formed"
 FT Region 87..158
 FT /label= B_chain
 XX
 PN MO9700945-AI.
 XX
 PD 09-JAN-1997.
 XX
 PF 21-JUN-1996; 96WO-US10669.
 XX

PR 23-JUN-1995; 95US-0000480.
 XX
 PA (UYHA-) UNIV HAMAIL.
 XX
 PI Chen H, Hu Z, Sun SSM, Xiong L;
 XX
 DR WPI; 1997-087372/08.
 XX
 PT Mabinlin nucleotide sequences - for prodn. of transgenic plants or
 PT recombinant proteins, useful as sweeteners, partic. in foods
 XX
 PS Disclosure; Fig 3; 38pp; English.
 XX
 CC The present sequence represents the mabinlin MBLIII. The encoding cDNA
 CC can modulate the expression of the gene encoding MBL in a cell or
 CC tissue. Mabinlin comes from the seeds of the Capparis masakakai plant,
 CC and is limited by its native availability, and so methods for producing
 CC recombinant mabinlin protein are being produced using chimeric genes
 CC such as MBLIII (see AAT60776). The products provide recombinant materials
 CC for the production of the MBL protein in practical amounts, and for the
 CC production of transgenic plants containing inherently sweet edible parts
 CC by virtue of production of MBL in situ. The MBL protein can be used to
 CC sweeten foods, beverages, animal feeds, chewing gum, dental hygiene
 CC products and pharmaceuticals.
 CC
 SQ Sequence 158 AA;
 AAW23588 Length: 158 September 12, 2002 12:06 Type: P Check: 8278 ..
 Found using 'consensus' (consens.key)
 1 MAKILLLTTLALFVLLANASIVRTVELEDEDNDENQPLCRQFOHQOVRACQRTLR
 42
 61 RRAQGLADELEDEVEDNEDENDQRGPARLRCNQLRQVDRPCVPLRQAHQOL
 121 YQGIIEGPROVRLFRARNLPNICKIPANGRCQPTRW
 153
 1 match found in sequence:
 aaw86264; Wild-type Brazil nut 2S albumin seed storage protein.
 (from "consensus-ags.pep")
 TOIG of: aaw86264 check: 7583 from: 1 to: 111
 ID AAW86264 standard; Protein: 111 AA.
 XX
 AC AAW86264;
 XX
 DT 01-MAR-1999 (first entry)
 XX
 DE Wild-type Brazil nut 2S albumin seed storage protein.
 XX
 KM Brazil Nut 2S albumin; seed storage protein; modified; BNCSS; BNL1;
 KM BNL1; BNL7; BNL8; BNL9; BNL13KW; ATZ51BN15; ATZ51BN13W;
 KM chimeric; dietary protein; essential amino acid.
 XX
 OS Bertholletia excelsa.
 XX
 PN WO9845458-AI.
 XX
 PD 15-OCT-1998.
 XX
 PE 06-APR-1998; 98WO-US06673.
 XX
 PR 08-APR-1997; 97US-0042827.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Gutteridge S;

```

XX WPI: 1998-609902/51.
DR N-PSDB; AAW80240.
XX
PT New modified Brazil Nut 2S albumin proteins - useful to provide
PT transgenic crops whose proteins have increased levels of
PT essential amino acids
XX
PS Example 1; Fig 1; 80pp; English.
XX
CC This represents a wild-type Brazil Nut 2S albumin seed storage
CC protein. The invention relates to modified Brazil Nut 2S albumin
CC proteins that are rich in essential amino acids and nucleic acids
CC encoding the modified proteins. The modified Brazil nut 2S albumin are
CC selected from a group comprising of BNCNSS, BN11, BN15, BN17, BN18, BN19,
CC BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes. Chimeric genes
CC comprising a nucleic acid fragment encoding a modified Brazil nut 2S
CC protein operably linked to suitable regulatory sequences, integrated into
CC the genome of a transformed plant can be used in methods for
CC increasing the essential amino acid content of seeds. The products can be
CC used to provide a human and animal dietary protein in seed crop plants
CC that is relatively high in essential amino acid content.
XX
SQ Sequence 111 AA:
AAW86264 Length: 111 September 12, 2002 12:06 Type: P Check: 7583 ..
Found using 'consensus' (consens.key)
1 MOECCREOMQOROMLSHCNRYMROQMESEPYOTMPRGMEPHMSCECQLEGMDESCRCE
5
61 GURMMRMROQEBMOPRGOMRRMRLAENIPSRCNLSFMRCPGSGIAGF
102
1 match found in sequence:
aaw86265 ; Modified Brazil nut 2S albumin gene BNCNSS amino acid sequence.
(from "consensus-agrs.pep")
TOIG of: aaw86265 check: 6078 from: 1 to: 111
ID AAW86265 standard; Protein; 111 AA.
XX
AC AAW86265;
XX
DT 01-MAR-1999 (first entry)
XX
DE Modified Brazil nut 2S albumin gene BNCNSS amino acid sequence.
XX
KW Brazil Nut 2S albumin; seed storage protein; modified; BNCNSS; BN11;
KW BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
KW chimeric; dietary protein; essential amino acid.
XX
OS Bertholletia excelsa.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 37 /label= R37K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 58 /label= R58K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 63 /note= "wild-type Arg is replaced with Lys"
FT /label= R63K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 82 /label= R82K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 83 /label= R83K
FT /note= "wild-type Arg is replaced with Lys"

```

```

FT Misc-difference 101 /label= R101K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 104 /label= F104M
FT /note= "wild-type Phe is replaced with Met"
FT Misc-difference /label= R37K
FT /note= "wild-type Arg is replaced with Lys"
XX
PN MO9845458-A1.
XX
XX 15-OCT-1998.
PD
XX 06-APR-1998; 98WO-US06673.
PF
XX 08-APR-1997; 97US-0042827.
PR
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA
XX Gutteridge S;
XX
DR WPI: 1998-609902/51.
DR N-PSDB; AAW80245.
XX
PT New modified Brazil Nut 2S albumin proteins - useful to provide
PT transgenic seed crops whose proteins have increased levels of
PT essential amino acids
XX
PS Claim 6; Fig 2; 80pp; English.
XX
CC The invention relates to modified Brazil Nut 2S albumin seed storage
CC proteins and nucleic acids encoding the modified proteins. The modified
CC Brazil nut 2S albumin proteins are rich in essential amino acids and the
CC encoding genes selected from a group comprising of BNCNSS, BN11, BN15,
CC BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
CC Chimeric genes comprising a nucleic acid fragment encoding a modified
CC Brazil nut 2S protein operably linked to suitable regulatory sequences,
CC integrated into the genome of a transformed plant can be used in
CC methods for increasing the essential amino acid content of seeds. The
CC products can be used to provide a human and animal dietary protein in
CC seed crop plants that is relatively high in essential amino acid content.
CC The present sequence represents the amino acid sequence of a modified
CC Brazil Nut 2S albumin gene BNCNSS.
XX
SQ Sequence 111 AA:
AAW86265 Length: 111 September 12, 2002 12:06 Type: P Check: 6078 ..
Found using 'consensus' (consens.key)
1 MOECCREOMQOROMLSHCNRYMROQMESEPYOTMPRGMEPHMSCECQLEGMDESCRCE
5
61 GURMMRMROQEBMOPRGOMRRMRLAENIPSRCNLSFMRCPGSGIAGF
102
1 match found in sequence:
aaw86266 ; Modified Brazil nut 2S albumin gene BN11 amino acid sequence.
(from "consensus-agrs.pep")
TOIG of: aaw86266 check: 5554 from: 1 to: 111
ID AAW86266 standard; Protein; 111 AA.
XX
AC AAW86266;
XX
DT 01-MAR-1999 (first entry)
XX
DE Modified Brazil nut 2S albumin gene BN11 amino acid sequence.
XX
KW Brazil Nut 2S albumin; seed storage protein; modified; BNCNSS; BN11;

```

KM BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
 KW chimeric; dietary protein; essential amino acid.
 XX
 OS Bertholletia excelsa.
 OS Synthetic.
 XX
 PN WO9845458-A1.
 XX
 PD 15-OCT-1998.
 XX
 PF 06-APR-1998; 98WO-US06673.
 XX
 PR 08-APR-1997; 97US-0042827.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Gutteridge S;
 XX
 DR WPI; 1998-609902/51.
 DR N-PSDB; AAV80246.
 XX
 PT New modified Brazil Nut 2S albumin proteins - useful to provide
 PT transgenic seed crops whose proteins have increased levels of
 PT essential amino acids
 XX
 PS Claim 6; Fig 3; 80pp; English.
 XX
 CC The invention relates to modified Brazil Nut 2S albumin seed storage
 CC proteins and nucleic acids encoding the modified proteins. The modified
 CC Brazil nut 2S albumin proteins are rich in essential amino acids and the
 CC encoding genes selected from a group comprising of BNCNS5, BN11, BN15,
 CC BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
 CC Chimeric genes comprising a nucleic acid fragment encoding a modified
 CC Brazil nut 2S protein operably linked to suitable regulatory sequences,
 CC integrated into the genome of a transformed plant can be used in
 CC methods for increasing the essential amino acid content of seeds. The
 CC products can be used to provide a human and animal dietary protein in
 CC seed crop plants that is relatively high in essential amino acid content.
 CC The present sequence represents the amino acid sequence of a modified
 CC Brazil Nut 2S albumin gene BN15. The modified BN15 has contains 11 Arg
 CC residues of the wild-type protein replaced with Lys residues.
 CC
 XX
 SQ Sequence 111 AA;
 AAW86266 Length: 111 September 12, 2002 12:06 Type: P Check: 5554 ..
 Found using 'consensus' (consens.key)

1 MOECCKEQMKOKMLSHCKMYWKQMESSPYQSMRKGMKMEPHMSBCCQLGEMDSCKCE
 5
 -----1-----
 61 GIKMMMMRQOEEMQPRGEOQMKMKMKLAENIPSKNLSPMKCPFGSGTAGF
 102

1 match found in sequence:
 aaw86267; Modified Brazil nut 2S albumin gene BN15 amino acid sequence.
 (from "consensus-ags.pep")
 TOIG of: aaw86267 check: 4826 from: 1 to: 111

ID AAW86267 standard; Protein; 111 AA.
 XX
 AC AAW86267;
 XX
 DT 01-MAR-1999 (first entry)
 XX
 DE Modified Brazil nut 2S albumin gene BN15 amino acid sequence.
 XX
 KW Brazil Nut 2S albumin; seed storage protein; modified; BNCNS5; BN11;
 KW BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
 KW chimeric; dietary protein; essential amino acid.
 OS
 XX

OS Bertholletia excelsa.
 OS Synthetic.
 XX
 PN WO9845458-A1.
 XX
 PD 15-OCT-1998.
 XX
 PF 06-APR-1998; 98WO-US06673.
 XX
 PR 08-APR-1997; 97US-0042827.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E I.
 XX
 PI Gutteridge S;
 XX
 DR WPI; 1998-609902/51.
 DR N-PSDB; AAV80247.
 XX
 PT New modified Brazil Nut 2S albumin proteins - useful to provide
 PT transgenic seed crops whose proteins have increased levels of
 PT essential amino acids
 XX
 PS Claim 6; Fig 4; 80pp; English.
 XX
 CC The invention relates to modified Brazil Nut 2S albumin seed storage
 CC proteins and nucleic acids encoding the modified proteins. The modified
 CC Brazil nut 2S albumin proteins are rich in essential amino acids and the
 CC encoding genes selected from a group comprising of BNCNS5, BN11, BN15,
 CC BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
 CC Chimeric genes comprising a nucleic acid fragment encoding a modified
 CC Brazil nut 2S protein operably linked to suitable regulatory sequences,
 CC integrated into the genome of a transformed plant can be used in
 CC methods for increasing the essential amino acid content of seeds. The
 CC products can be used to provide a human and animal dietary protein in
 CC seed crop plants that is relatively high in essential amino acid content.
 CC The present sequence represents the amino acid sequence of a modified
 CC Brazil Nut 2S albumin gene BN15. The modified BN15 has contains 15 Arg
 CC residues of the wild-type protein replaced with Lys residues.
 CC
 XX
 SQ Sequence 111 AA;
 AAW86267 Length: 111 September 12, 2002 12:06 Type: P Check: 4826 ..
 Found using 'consensus' (consens.key)

1 MOECCKEQMKOKMLSHCKMYWKQMESSPYQSMRKGMKMEPHMSBCCQLGEMDSCKCE
 5
 -----1-----
 61 GIKMMMMRQOEEMQPRGEOQMKMKMKLAENIPSKNLSPMKCPFGSGTAGF
 102

1 match found in sequence:
 aaw86268; Modified Brazil nut 2S albumin gene BN17 amino acid sequence.
 (from "consensus-ags.pep")
 TOIG of: aaw86268 check: 4618 from: 1 to: 111

ID AAW86268 standard; Protein; 111 AA.
 XX
 AC AAW86268;
 XX
 DT 01-MAR-1999 (first entry)
 XX
 DE Modified Brazil nut 2S albumin gene BN17 amino acid sequence.
 XX
 KW Brazil Nut 2S albumin; seed storage protein; modified; BNCNS5; BN11;
 KW BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
 KW chimeric; dietary protein; essential amino acid.
 OS Bertholletia excelsa.
 OS Synthetic.
 XX

```

PN WO9845458-A1.
XX
XX 15-OCT-1998.
PD
XX 06-APR-1998; 98WO-US06673.
PE
XX 08-APR-1997; 97US-0042827.
PR
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA
XX Gutteridge S;
PI
XX WPI; 1998-609902/51.
DR
XX N-PSDB; AAV80248.
DR
XX
XX New modified Brazil Nut 2S albumin proteins - useful to provide
PT transgenic seed crops whose proteins have increased levels of
PT essential amino acids
XX
XX Claim 6; Fig 5; 80pp; English.
PS
XX The invention relates to modified Brazil Nut 2S albumin seed storage
XX proteins and nucleic acids encoding the modified proteins. The modified
CC Brazil nut 2S albumin proteins are rich in essential amino acids and the
CC encoding genes selected from a group comprising of BNCNS5, BN11, BN15,
CC BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
CC Chimeric genes comprising a nucleic acid fragment encoding a modified
CC Brazil nut 2S protein operably linked to suitable regulatory sequences,
CC integrated into the genome of a transformed plant can be used in
CC methods for increasing the essential amino acid content of seeds. The
CC products can be used to provide a human and animal dietary protein in
CC seed crop plants that is relatively high in essential amino acid content.
CC The present sequence represents the amino acid sequence of a modified
CC Brazil Nut 2S albumin gene BN17. The modified BN17 has all the 15 Arg
CC residues, Gly at position 105 and Ser at position 107 of the wild-type
CC protein replaced with Lys residues
XX
XX Sequence 111 AA;
SO
AAW86268 Length: 111 September 12, 2002 12:06 Type: P Check: 4618 ..
Found using 'consensus' (consens.key)
1
-----|-----
1 MOEBCKEQMOKOKMLSHCKMTMKQOMESPYQSMPPKGMPEHMCCEQBLEGMDSCCKCE
5
GLKMMKMMQOEEMQPKGEQMKKKMKLAENIPSKCNLSPMKCPFKGIAGF
61 102
-----|-----
1 match found in sequence:
aaw86269 ; Modified Brazil nut 2S albumin gene BN18 amino acid sequence.
(from "consensus-ags.pep")
TOIG of: aaw86269 check: 4266 from: 1 to: 111
ID AAW86269 standard; Protein: 111 AA.
XX
XX AAW86269;
AC
XX 01-MAR-1999 (first entry)
DT
XX Modified Brazil nut 2S albumin gene BN18 amino acid sequence.
DE
XX
XX Brazil Nut 2S albumin; seed storage protein; modified; BNCNS5; BN11;
KW BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
KW chimeric; dietary protein; essential amino acid.
XX
XX Bertholletia excelsa.
OS Synthetic.
XX
XX WO9845458-A1.
XX

```

```

PD 15-OCT-1998.
XX
XX 06-APR-1998; 98WO-US06673.
PE
XX 08-APR-1997; 97US-0042827.
PR
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA
XX Gutteridge S;
PI
XX WPI; 1998-609902/51.
DR
XX N-PSDB; AAV80249.
DR
XX
XX New modified Brazil Nut 2S albumin proteins - useful to provide
PT transgenic seed crops whose proteins have increased levels of
PT essential amino acids
XX
XX Claim 6; Fig 6; 80pp; English.
PS
XX The invention relates to modified Brazil Nut 2S albumin seed storage
XX proteins and nucleic acids encoding the modified proteins. The modified
CC Brazil nut 2S albumin proteins are rich in essential amino acids and the
CC encoding genes selected from a group comprising of BNCNS5, BN11, BN15,
CC BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
CC Chimeric genes comprising a nucleic acid fragment encoding a modified
CC Brazil nut 2S protein operably linked to suitable regulatory sequences,
CC integrated into the genome of a transformed plant can be used in
CC methods for increasing the essential amino acid content of seeds. The
CC products can be used to provide a human and animal dietary protein in
CC seed crop plants that is relatively high in essential amino acid content.
CC The present sequence represents the amino acid sequence of a modified
CC Brazil Nut 2S albumin gene BN18. The modified BN18 has all the 15 Arg
CC residues, Gly at position 105, and Ser at positions 44 and 107 of the
CC wild-type protein replaced with Lys residues
XX
XX Sequence 111 AA;
SO
AAW86269 Length: 111 September 12, 2002 12:06 Type: P Check: 4266 ..
Found using 'consensus' (consens.key)
1
-----|-----
1 MOEBCKEQMOKOKMLSHCKMTMKQOMESPYQSMPPKGMPEHMCCEQBLEGMDSCCKCE
5
GLKMMKMMQOEEMQPKGEQMKKKMKLAENIPSKCNLSPMKCPFKGIAGF
61 102
-----|-----
1 match found in sequence:
aaw86270 ; Modified Brazil nut 2S albumin gene BN19 amino acid sequence.
(from "consensus-ags.pep")
TOIG of: aaw86270 check: 5234 from: 1 to: 111
ID AAW86270 standard; Protein: 111 AA.
XX
XX AAW86270;
AC
XX 01-MAR-1999 (first entry)
DT
XX Modified Brazil nut 2S albumin gene BN19 amino acid sequence.
DE
XX
XX Brazil Nut 2S albumin; seed storage protein; modified; BNCNS5; BN11;
KW BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
KW chimeric; dietary protein; essential amino acid.
XX
XX Bertholletia excelsa.
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Misc-difference 4 /label= E28K
FT /note= "wild-type Glu is replaced with Lys"
FT

```


FT	Misc-difference	11	/label= R11K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	16	/label= R16K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	19	/label= R19K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	23	/label= R23K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	27	/label= E27M
FT	/note= "wild-type Glu is replaced with Met"		
FT	Misc-difference	28	/label= E28K
FT	/note= "wild-type Glu is replaced with Lys"		
FT	Misc-difference	36	/label= R36K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	37	/label= R37K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	58	/label= R58K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	63	/label= R63K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	68	/label= R68K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	77	/label= R77K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	82	/label= R82K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	83	/label= R83K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	86	/label= R86K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	94	/label= R94K
FT	/note= "wild-type Arg is replaced with Lys"		
FT	Misc-difference	101	/label= R101K
FT	/note= "wild-type Arg is replaced with Lys"		
PD	W0845458-A1.		
PD	15-OCT-1998.		
PE	06-APR-1998;	98WC-US06673.	
PR	08-APR-1997;	97US-0042827.	
PA	(DUPO) DU PONT DE NEMOURS & CO E. I.		
PI	Gutteridge S;		
DR	WPI; 1998-609902/51.		
DR	N-PSDB; AAV80250.		
PT	New modified Brazil Nut 2S albumin proteins - useful to provide		
PT	transgenic seed crops whose proteins have increased levels of		
PT	essential amino acids		
PS	Claim 6; Fig 7; 80pp; English.		

```

CC The invention relates to modified Brazil Nut 2S albumin seed storage
CC proteins and nucleic acids encoding the modified proteins. The modified
CC Brazil nut 2S albumin proteins are rich in essential amino acids and the
CC encoding genes selected from a group comprising of BNCNS, BN11, BN15,
CC BN17, BN18, BN19, BN153KW, AT251BN15, AT251BN153W gene.
CC Chimeric genes comprising a nucleic acid fragment encoding a modified
CC Brazil nut 2S protein operably linked to suitable regulatory sequences,
CC integrated into the genome of a transformed plant can be used in
CC methods for increasing the essential amino acid content of seeds. The
CC products can be used to provide a human and animal dietary protein in
CC seed crop plants that is relatively high in essential amino acid content.
CC The present sequence represents the amino acid sequence of a modified
CC Brazil Nut 2S albumin gene BN19.
SQ
XX
SQ Sequence 111 AA:
AAM6270 Length: 111 September 12, 2002 12:06 Type: P Check: 5234 ..
Found using 'consensus' (consens.key)

1-----|
1 MOECKEOMOKRMKLSHCKMYMKQOMMKSPYQSMFKKGMPEHMSCECOLEGMDESCKE
5
-----|
GLKMMMKMKQOEEMQPKGEOMKKMMKLAENIPSKCNLSPMKCPGGSGIAGF

```

1 match found in sequence:	
aa86271	Modified Brazil1 nut 2S albumin gene BN153KW amino acid sequence (from "consensus-eggs.pep")
TOIG of: aa86271 Check: 5147 from: 1 to: 111	
ID	AAW86271 standard; Protein; 111 AA.
XX	
AC	AAW86271;
XX	
DT	01-MAR-1999 (first entry)
XX	
DE	Modified Brazil1 nut 2S albumin gene BN153KW amino acid sequence.
XX	
KW	Brazil1 Nut 2S albumin; seed storage protein; modified; BNCNS; BN11;
KW	BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
KM	chimeric; dietary protein; essential amino acid.
XX	
OS	Bertholletia excelsa.
OS	Synthetic.
XX	
FH	Key
FT	Misc-difference 4 Location/Qualifiers
FT	/label=EK
FT	/note="wild-type Glu is replaced with Lys"
FT	Misc-difference 6
FT	/label=R6K
FT	/note="wild-type Arg is replaced with Lys"
FT	Misc-difference 11
FT	/label=R1K
FT	/note="wild-type Arg is replaced with Lys"
FT	Misc-difference 16
FT	/label=S16K
FT	/note="wild-type Ser is replaced with Lys"
FT	Misc-difference 19
FT	/label=R19K
FT	/note="wild-type Arg is replaced with Lys"
FT	Misc-difference 23
FT	/label=R23K
FT	/note="wild-type Arg is replaced with Lys"
FT	Misc-difference 36
FT	/label=R36K
FT	/note="wild-type Arg is replaced with Lys"
FT	Misc-difference 37
FT	/label=R37K
FT	/note="wild-type Arg is replaced with Lys"

```

FT Misc-difference 41
FT /label= P41K
FT /note= "wild-type Pro is replaced with Lys"
FT Misc-difference 42
FT /label= W42H
FT /note= "wild-type Trp is replaced with His"
FT Misc-difference 58
FT /label= R58K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 63
FT /label= R63K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 68
FT /label= R68K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 77
FT /label= R77K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 82
FT /label= R82K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 83
FT /label= R83K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 86
FT /label= R86K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 94
FT /label= R94K
FT /note= "wild-type Arg is replaced with Lys"
FT Misc-difference 101
FT /label= R101K
FT /note= "wild-type Arg is replaced with Lys"
FT W09845458-A1.
FT 15-OCT-1998.
FT 06-APR-1998; 98WO-US06673.
FT 08-APR-1997; 97US-0042827.
FT (DUPO ) DU PONT DE NEMOURS & CO E. I.
FT Gutteridge S;
FT WPI: 1998-609902/51.
FT N-PSDB; AAV80251.
FT New modified Brazil Nut 2S albumin proteins - useful to provide
FT transgenic seed crops whose proteins have increased levels of
FT essential amino acids
FT Claim 6; Fig 8; 80pp; English.
XX The invention relates to modified Brazil Nut 2S albumin seed storage
XX proteins and nucleic acids encoding the modified proteins. The modified
XX Brazil nut 2S albumin proteins are rich in essential amino acids and the
XX encoding genes selected from a group comprising of BNCNS, BN11, BN15,
XX BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
XX Chimeric genes comprising a nucleic acid fragment encoding a modified
XX Brazil nut 2S protein operably linked to suitable regulatory sequences,
XX integrated into the genome of a transformed plant can be used in
XX methods for increasing the essential amino acid content of seeds. The
XX products can be used to provide a human and animal dietary protein in
XX seed crop plants that is relatively high in essential amino acid content.
XX The present sequence represents the amino acid sequence of a modified
XX Brazil Nut 2S albumin gene BN153KW.
XX
XX Sequence 111 AA;
AAW6271 Length: 111 September 12, 2002 12:06 Type: P Check: 5147 ..
Found using 'consensus' (consens.key)

```

```

1 MOEKCEQOMOKOKMLKHCCKMYNKQMEESPYSMPKGMKMEKMSBCCBQLGMEBSCKE
5
61 GIKMMMKQOEKQPKGEQMKMKMLAENIPSKCNLSPMKCPGGSIAGF
102
-----
1 match found in sequence:
aaw86272; Modified Brazil nut 2S albumin gene AT2S1BN15 amino acid sequence.
(from "consensus_ags.pep")
TOIG of: aaw86272 check: 4544 from: 1 to: 149
ID AAW86272 standard; Protein; 149 AA.
XX
XX AAW86272;
AC
XX
XX 01-MAR-1999 (first entry)
DE
XX
XX Modified Brazil nut 2S albumin gene AT2S1BN15 amino acid sequence.
XX
XX Brazil Nut 2S albumin; seed storage protein; modified; BNCNS; BN11;
XX BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
XX chimeric; dietary protein; essential amino acid.
XX
XX Bertholletia excelsa.
OS
XX Synthetic.
XX
XX W09845458-A1.
XX 15-OCT-1998.
XX 06-APR-1998; 98WO-US06673.
XX 08-APR-1997; 97US-0042827.
XX (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX
XX Gutteridge S;
XX WPI: 1998-609902/51.
XX N-PSDB; AAV80252.
XX New modified Brazil Nut 2S albumin proteins - useful to provide
XX transgenic seed crops whose proteins have increased levels of
XX essential amino acids
XX Claim 6; Fig 10; 80pp; English.
XX
XX The invention relates to modified Brazil Nut 2S albumin seed storage
XX proteins and nucleic acids encoding the modified proteins. The modified
XX Brazil nut 2S albumin proteins are rich in essential amino acids and the
XX encoding genes selected from a group comprising of BNCNS, BN11, BN15,
XX BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
XX Chimeric genes comprising a nucleic acid fragment encoding a modified
XX Brazil nut 2S protein operably linked to suitable regulatory sequences,
XX integrated into the genome of a transformed plant can be used in
XX methods for increasing the essential amino acid content of seeds. The
XX products can be used to provide a human and animal dietary protein in
XX seed crop plants that is relatively high in essential amino acid content.
XX The present sequence represents the amino acid sequence of a modified
XX Brazil Nut 2S albumin gene AT2S1BN15.
XX
XX Sequence 149 AA;
AAW6272 Length: 149 September 12, 2002 12:06 Type: P Check: 4544 ..
Found using 'consensus' (consens.key)

```

61 KOQMEESPYSMPKKGMPHMSCECCOLEGMDSCKEGLKMMAMKMQEEMOKRGEOMK

121 KMMKLAENIPSKNLSPMKCPFGSGIAGF
140

1 match found in sequence:
aaw86273 ; Modified Brazil nut 2S albumin gene AT2S1BN19 amino acid sequence.
(from "consensus.ags.pep")
T0IG of: aaw86273 check: 4914 from: 1 to: 149

ID AAW86273 standard; Protein: 149 AA.
XX
AC AAW86273;
XX
DT 01-MAR-1999 (first entry)
XX
DE Modified Brazil nut 2S albumin gene AT2S1BN19 amino acid sequence.
XX
KW Brazil Nut 2S albumin; seed storage protein; modified; BNCNSG; BN11;
KW BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
KW chimeric; dietary protein; essential amino acid.
XX
OS Bertholletia excelsa.
OS Synthetic.
XX
PN MO9845458-A1.
XX
PD 15-OCT-1998.
XX
PE 06-APR-1998; 98WO-US06673.
XX
PR 08-APR-1997; 97US-0042827.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Gutteridge S;
XX
DR WPI: 1998-609902/51.
DR N-PSDB: AAV80273.
XX
PT New modified Brazil Nut 2S albumin proteins - useful to provide
PT transgenic seed crops whose proteins have increased levels of
PT essential amino acids
XX
PS Claim 6; Fig 11; 80pp; English.
XX
CC The invention relates to modified Brazil Nut 2S albumin seed storage
CC proteins and nucleic acids encoding the modified proteins. The modified
CC Brazil nut 2S albumin proteins are rich in essential amino acids and the
CC encoding genes selected from a group comprising of BNCNSG, BN11, BN15,
CC BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
CC Chimeric genes comprising a nucleic acid fragment encoding a modified
CC Brazil nut 2S protein operably linked to suitable regulatory sequences,
CC integrated into the genome of a transformed plant can be used in
CC methods for increasing the essential amino acid content of seeds. The
CC products can be used to provide a human and animal dietary protein in
CC seed crop plants that is relatively high in essential amino acid content.
CC The present sequence represents the amino acid sequence of a modified
CC Brazil Nut 2S albumin gene AT2S1BN19.
XX
SQ Sequence 149 AA;

AAW86273 Length: 149 September 12, 2002 12:06 Type: P Check: 4914 ..
Found using 'consensus' (consens.key)

1 MANKFLVCAALALCFLLTNASIVRTVVEEEDDATNHMOECCKEQOMOKMISHCKMYM
43

61 KOQMEESPYSMPKKGMPHMSCECCOLEGMDSCKEGLKMMAMKMQEEMOKRGEOMK

121 KMMKLAENIPSKNLSPMKCPFGSGIAGF
140

1 match found in sequence:
aaw86274 ; Modified Brazil nut 2S albumin gene AT2S1BN153W amino acid sequence
(from "consensus.ags.pep")
T0IG of: aaw86274 check: 5074 from: 1 to: 149

ID AAW86274 standard; Protein: 149 AA.
XX
AC AAW86274;
XX
DT 01-MAR-1999 (first entry)
XX
DE Modified Brazil nut 2S albumin gene AT2S1BN153W amino acid sequence.
XX
KW Brazil Nut 2S albumin; seed storage protein; modified; BNCNSG; BN11;
KW BN15; BN17; BN18; BN19; BN153KW; AT2S1BN15; AT2S1BN19; AT2S1BN153W;
KW chimeric; dietary protein; essential amino acid.
XX
OS Bertholletia excelsa.
OS Synthetic.
XX
PN MO9845458-A1.
XX
PD 15-OCT-1998.
XX
PE 06-APR-1998; 98WO-US06673.
XX
PR 08-APR-1997; 97US-0042827.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Gutteridge S;
XX
DR WPI: 1998-609902/51.
DR N-PSDB: AAV80254.
XX
PT New modified Brazil Nut 2S albumin proteins - useful to provide
PT transgenic seed crops whose proteins have increased levels of
PT essential amino acids
XX
PS Claim 6; Fig 12; 80pp; English.
XX
CC The invention relates to modified Brazil Nut 2S albumin seed storage
CC proteins and nucleic acids encoding the modified proteins. The modified
CC Brazil nut 2S albumin proteins are rich in essential amino acids and the
CC encoding genes selected from a group comprising of BNCNSG, BN11, BN15,
CC BN17, BN18, BN19, BN153KW, AT2S1BN15, AT2S1BN19, AT2S1BN153W genes.
CC Chimeric genes comprising a nucleic acid fragment encoding a modified
CC Brazil nut 2S protein operably linked to suitable regulatory sequences,
CC integrated into the genome of a transformed plant can be used in
CC methods for increasing the essential amino acid content of seeds. The
CC products can be used to provide a human and animal dietary protein in
CC seed crop plants that is relatively high in essential amino acid content.
CC The present sequence represents the amino acid sequence of a modified
CC Brazil Nut 2S albumin gene AT2S1BN153W.
XX
SQ Sequence 149 AA;

AAW86274 Length: 149 September 12, 2002 12:06 Type: P Check: 5074 ..
Found using 'consensus' (consens.key)

1 MANKFLVCAALALCFLLTNASIVRTVVEEEDDATNHMOECCKEQOMOKMISHCKMYM
43

61 KOOMESPYSOMPKMEPHMSECCBOLBGMDESCKEGLKMMMKMOJEMOPKGEOMK

121 KMMKLANIPWKCINLSPMKCPWGKLAGF 140

1 match found in sequence:
aay54571; Amino acid sequence of the CM17 protein of wheat.
(from "consensus_ags.pep")
TOIG of: aay54571 check: 7626 from: 1 to: 123

ID AAY54571 standard; Protein; 123 AA.

AC AAY54571;

DT 25-APR-2000 (first entry)

DE Amino acid sequence of the CM17 protein of wheat.

KM CM17; glutenin; seed storage protein; gluten; bread; pasta; noodle;
KW breakfast cereal; snack food; cake; pastry; flour based sauce; film;
KM coating; adhesive; building material; packaging material; grain;
KW lipid binding domain.

OS Triticum aestivum.

PN WO200002914-A1.

PD 20-JAN-2000.

PE 12-JUL-1999; 99WO-AU00563.

PR 10-JUL-1998; 98AU-0004604.

PA (CSIR) COMMONWEALTH SCI & IND RES ORG.

PA (GOOD-) GOODMAN FIELDER LTD.

PA (LIMA-) GRP LIMAGRAIN PACIFIC PTY LTD.

PI Appels R, Morell M, Bekes F, Tamas L;

DR WPI; 2000-147597/13.

PT Modifying glutenin or seed-storage protein for preparing foodstuffs,
PT films, coatings, packing materials, adhesives and building materials -
PS Disclosure; Page 73; 76pp; English.

CC The present sequence represents the CM17 protein of wheat. The lipid
CC binding domain of the CM17 gene was used to modify glutenins and seed
CC storage proteins, in the method of the invention. The specification
CC describes a method for producing a modified glutenin or seed storage
CC protein, by adding to the protein a domain that confers the ability
CC to incorporate into gluten, or to bind a ligand or other macromolecule.
CC The domain can be any domain that will bind ligands that may be useful
CC in food preparation or in food composition, e.g. a domain that binds
CC lipids or starch. The method is used for producing modified glutenins.
CC glutenins and seed-storage protein are useful for preparing food
CC products such as leavened or unleavened breads, pasta, noodles,
CC breakfast cereals, snack foods, cakes, pastries and foods containing
CC flour based sauces. Glutenins and seed-storage protein are also useful
CC for preparing nonfood products such as films, coatings, adhesives,
CC building materials and packaging materials. Grain or parts of grain
CC containing the modified glutenin and seed-storage protein is useful for
CC preparing food products. The modified glutenins and seed-storage proteins
CC are also useful as modifiers of food products in food industry.

SQ Sequence 123 AA;

AAY54571 Length: 123 September 12, 2002 12:06 Type: P Check: 7626

Found using 'consensus' (consens. key)

1 AIGNEOCTPWTSTLITPLPSCRYVEBOACRIEMGPPIYIAKOECCOJLANIPQOCRCOA

61 LRYFMGPKSRPDOSGLMELPGCPREVOMNEPILVTPGYCNLTTHNTPYCLAMESSOMS 111

121 CIM

1 match found in sequence:
aay58049; Puroindoline B protein sequence.
(from "consensus_ags.pep")
TOIG of: aay58049 check: 8339 from: 1 to: 148

ID AAY58049 standard; Protein; 148 AA.

AC AAY58049;

DT 14-MAR-2000 (first entry).

DE Puroindoline B protein sequence.

KM Puroindoline B; friabilin; grain softness; wheat; starch; grain hardness;
KW grain texture; transgenic plant; sorghum; rice; barley; maize; rye;
KM oat; triticale.

OS Triticum aestivum.

PN WO961580-A2.

PD 02-DEC-1999.

PE 21-MAY-1999; 99WO-US11379.

PR 22-MAY-1998; 98US-0083852.

PA (RERE-) RES & DEV INST INC.

PA (USDA) US DEPT OF AGRICULTURE.

PI Morris CF, Giroux MJ;

DR WPI; 2000-062702/05.

PT Altering grain hardness in plants by modifying expression of
PT puroindoline proteins -
PS Examples; Page 62; 75pp; English.

CC This is the wheat puroindoline B protein sequence. Friabilin is a 15kD
CC marker protein for grain softness, and is composed of Puroindoline A and
CC Puroindoline B. Friabilin is present in high amounts on the surface of
CC soft wheat starch, and in low amounts on hard wheat starch. The
CC puroindolines have tryptophan-rich, hydrophobic domains which have
CC affinity for binding lipids. Puroindoline A and puroindoline B are used
CC in the production of transgenic plants with modified grain hardness. The
CC invention can be used to produce plants with softer grain texture, where
CC at least one parent plant has a hard texture grain. The softer grained
CC plants are produced through the introduction of a nucleic acid sequence
CC which operably encodes a puroindoline protein into a cell from the parent
CC plant, progeny plants are generated for the cell containing the nucleic
CC acid sequence. The methods can be used for producing transgenic plants
CC with softer textured grain from plants such as durum wheat, sorghum,
CC rice, barley and maize. They can also be used for producing transgenic
CC plants with harder textured grain from plants such as wheat, rye,
CC triticale and oat. The modification of the grain texture results in
CC softer grains which require less milling.

SQ Sequence 148 AA;

AAV58049 Length: 148 September 12, 2002 12:06 Type: P Check: 8339 ..
Found using 'consensus' (consens.key)

1 MKTFLALIALVASTTFAQYSEVGWYNEVGSGSQCCPOERPKLSSCKDYMERCT
40

61 MKDFVTWPTKWKKGCEHEVREKCKKQLSQIAPQCRCDISIRVIGRLGFLGIMRGV

121 FKQLRAQSLPSKCMGADCKPPSGYYW
140

1 match found in sequence:
aay58050 : Serine substituted puroindoline B protein sequence.
(from "consensus.ags.pep")
TOIG of: aay58050 check: 8555 from: 1 to: 148

ID AAY58050 standard; Protein; 148 AA.
XX AAY58050;
AC
XX
XX
DT 14-MAR-2000 (first entry)
XX
XX Serine substituted puroindoline B protein sequence.
DE
XX Puroindoline B; friablin; grain softness; wheat; starch; grain hardness;
KW grain texture; transgenic plant; sorghum; rice; barley; maize; rye;
KM oat; triticale.
XX
OS Triticum aestivum.
XX
XX WO9961580-A2.
XX
XX 02-DEC-1999.
XX
XX 21-MAY-1999; 99WO-US11379.
XX
XX 22-MAY-1998; 98US-0083852.
XX
XX (RERE-) RES & DEV INST. INC.
PA (USDA) US DEPT OF AGRICULTURE.
XX
XX
XX Morris CF, Giroux MJ;
PI
XX
XX
DR WPI: 2000-062702/05.
DR N-PSDB: AA247521.
XX
XX

Altering grain hardness in plants by modifying expression of
puroindoline proteins -
Disclosure: Page 63; 75pp; English.

CC This is the wheat serine substituted puroindoline B protein sequence.
CC Friablin is a 15kD marker protein for grain softness, and is composed of
CC puroindoline A and puroindoline B. Friablin is present in high amounts
CC on the surface of soft wheat starch, and in low amounts on hard wheat
CC starch. The puroindolines have tryptophan-rich, hydrophobic domains which
CC have affinity for binding lipids. Puroindoline A and puroindoline B are
CC used in the production of transgenic plants with modified grain hardness.
CC The invention can be used to produce plants with softer grain texture,
CC where at least one parent plant has a hard texture grain. The softer
CC grained plants are produced through the introduction of a nucleic acid
CC sequence which operably encodes a puroindoline protein into a cell from
CC the parent plant; progeny plants are generated for the cell containing
CC the nucleic acid sequence. The methods can be used for producing
CC transgenic plants with softer textured grain from plants such as durum
CC wheat, sorghum, rice, barley and maize. They can also be used for
CC producing transgenic plants with harder textured grain from plants such

CC as wheat, rye, triticale and oat. The modification of the grain texture
CC results in softer grains which require less milling.
XX
XX SQ Sequence 148 AA;
AAV58050 Length: 148 September 12, 2002 12:06 Type: P Check: 8555 ..
Found using 'consensus' (consens.key)

1 MKTFLALIALVASTTFAQYSEVGWYNEVGSGSQCCPOERPKLSSCKDYMERCT
40

61 MKDFVTWPTKWKKGCEHEVREKCKKQLSQIAPQCRCDISIRVIGRLGFLGIMRGV

121 FKQLRAQSLPSKCMGADCKPPSGYYW
140

-- Search Statistics --

Times:	CPU	Total Elapsed
	00:00:00.00	00:00:01.00
Number of sequences searched:		22
Number of sequence hits:		22
Number of separate matches:		22
Number of sequence hits saved:		0

THIS PAGE BLANK (USPTO)

> O <
0110 Intelligence
> O <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "consensus-is" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.

Selected sequence key from "consens.key":

consensus (AA) ID consensus AA preliminary pattern
followed by

1 C any character from 9 to 13 times
2 C
2 C
2 q or e or r or k
2 any character
2 a or f or g or i or l or m or p or t or v or w or y
2 l or i or v or m
2 any character from 15 to 39 times
2 CC
2 C or d or e or h or k or n or q or r or s
2 q or e or h
2 l or v
2 any character
2 any character
2 any character
2 any character
2 any character
2 C
2 any character
2 C
2 any character
2 l or i
2 any character from 13 to 56 times
2 g
2 any character from 15 to 26 times
2 C
2 any character
2 v or i or l or m
2 any character from 1 to 8 times
2 C

Selected data banks and files:

Data bank : Issued_AA , all entries

-- Output Parameters --

Format Options: File Options:
Nucleic acid code matching Exact Indirect file
Find non-matching hits only No Sequence or key file
Report key used Yes List of hits
Note position of hit Yes Hit display
Display full annotations Yes Name and annotations
Sequence context 50 Yes

-- Run Parameters --

Run mode Batch
Time to start comparison now
Notify at end of run No

1 match found in sequence:
US-08-670-186-2 ; Sequence 2, Application US/08670186
(from "/srch/1aa/5B_COMB.pep")
Sequence 2, Application US/08670186

Patent No. 5859343

GENERAL INFORMATION:

APPLICANT: SUN, SAMUEL S.M.
APPLICANT: XIONG, LIWEN
APPLICANT: HU, ZHONG
APPLICANT: CHEN, HANG

TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
CITY: WASHINGTON
STATE: DC

COUNTRY: USA

ZIP: 20006-1888

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/670,186
FILING DATE: 21-JUN-1996

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 23461-20007.00
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168

TELEX: 90-4030 MRSNFOERSMSH
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 155 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

Found using 'consensus' (consens.key)

1 MAKILFLFATLALFVLNANSIQTTVEVDEEDNQLMRCORFLQRLRACQRFHRR

40

61 AQFGQPEDEVEDDNDENQRPALRQCCNQLRQVPCVPLRQAQVLAQRI

149

121 IQGPQRLRLFDARNPINICINIGACPRAMP

149

1 match found in sequence:
US-08-670-186-4 ; Sequence 4, Application US/08670186
(from "/srch/1aa/5B_COMB.pep")
Sequence 4, Application US/08670186
Patent No. 5859343

GENERAL INFORMATION:

APPLICANT: SUN, SAMUEL S.M.
APPLICANT: XIONG, LIWEN
APPLICANT: HU, ZHONG

APPLICANT: CHEN, HANG

TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
CITY: WASHINGTON
STATE: DC

COUNTRY: USA

ZIP: 20006-1888
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/670,186
 FILING DATE: 21-JUN-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 23461-20007.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-1500
 TELEX: 90-4030 MRSNFOERSWSH
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 158 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 Found using 'consensus' (consens.key)

1 MAKILLITLALFVLLANASIRRTVELDEEDNDENQPLCRPQOQHOLRACQRYLR
 42
 61 RRAQSGGLADELEDEVEDENEDENQGRPALRLCCNQLROYNKPCCVPLRQAHHQOL

121 YQGIISGPRQVRRLFAARLPLNICKIPAVGRQCFTRW
 153

1 match found in sequence:
 US-08-670-186-6; Sequence 6, Application US/08670186
 (from "/srch/1aa/58.COMB.pep")
 Sequence 6, Application US/08670186
 Patent No. 5859343
 GENERAL INFORMATION:
 APPLICANT: SUN, SAMUEL S.M.
 APPLICANT: XIIONG, LIWEN
 APPLICANT: HU, ZHONG
 APPLICANT: CHEN, HANG
 TITLE OF INVENTION: RECOMBINANT SWEET PROTEIN MABINLIN
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 PENNSYLVANIA AVE NW, STE. 5500
 CITY: WASHINGTON
 STATE: DC
 COUNTRY: USA
 ZIP: 20006-1888
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/670,186
 FILING DATE: 21-JUN-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 23461-20007.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 882-0168
 TELEX: 90-4030 MRSNFOERSWSH

INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 158 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 Found using 'consensus' (consens.key)

1 MAKILLITLALFVLLANASIRRTVELDEEDNDENQPLCRPQOQHOLRACQRYLR
 42
 61 RRAQSGGLADELEDEVEDENEDENQGRPALRLCCNQLROYNKPCCVPLRQAHHQOL
 121 YQGIISGPRQVRRLFAARLPLNICKIPAVGRQCFTRW
 153

-- Search Statistics --
 Times: CPU Total Elapsed
 00:01:13.10 00:01:14.00
 Number of sequences searched: 23168
 Number of sequence hits: 3
 Number of separate matches: 3
 Number of sequence hits saved: 0

181 FAAGQY

1 match found in sequence:
JC5379 : TOIG of: JC5379 check: 7604 from: 1 to: 155
(from "consensus_pir.pep")
TOIG of: JC5379 check: 7604 from: 1 to: 155

P1:JC5379 - mablinin II precursor - Yunnan caper
C:Species: Capparis masaiakal (Yunnan caper)
C:Date: 02-Jun-1997 #sequence_revision 12-Sep-1997 #text_change 20-Jun-2000
C:Accession: JC5379, PC4316
R:Nirasawa, S.; Masuda, Y.; Nakaya, K.; Kurihara, Y.
Gene 181, 225-227, 1996
A:Title: Cloning and sequencing of a cDNA encoding a heat-stable sweet protein, mablinin II.
A:Reference number: JC5379; MUID:97128796

A:Accession: JC5379
A:Molecule type: mRNA
A:Residues: 1-155 <NIR1>
A:Cross-references: DDBJ:D83997; NID:g1817545; PIDN:BAAL2204.1; PID:g1817546
A:Accession: PC4316
A:Molecule type: Protein
A:Residues: 36-41,149-154 <NIR2>
A:Experimental source: seed
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: sweet taste

F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-35/Domain: amino-terminal propeptide #status predicted <PRO>
F:36-68/Domain: mablinin II #status experimental <MAI>
F:69-83/Domain: mablinin II A chain #status experimental <CHA>
F:83-154/Domain: mablinin II B chain #status experimental <CHB>

JC5379 length: 155 September 12, 2002 12:08 Type: P Check: 7604
Found using 'consensus' (consens.key)

1 MAKLFPLFATLALFLLANASIQTTVIEVDEEDNQLRCCROFLQHLRACQRFTHRR
40

61 AOGGPGDELEDEVEDNDNDENQRPALRCCNQLRQVDRVCVPLRQAAGQVLOROI

121 IQGPGQLRLFLDARNLPLNICNPINIGTCPRTPP
149

1 match found in sequence:
nmw2 : TOIG of: nmw2 check: 3182 from: 1 to: 170
(from "consensus_pir.pep")
TOIG of: nmw2 check: 3182 from: 1 to: 170

P1:NMW2 - 2S albumin 2 precursor - Arabidopsis thaliana
N:Alternate names: seed storage protein AT252
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 03-Dec-1999
C:Accession: JA0162, S34677; T06045
R:Krebers, E.; Herdies, L.; de Clercq, A.; Seurinck, J.; Leemans, J.; Van Damme, J.; Segura, M.; Gheysen, G.; Van Montagu, M.; Vandekerckhove, J.
Plant Physiol. 87, 859-866, 1988
A:Title: Determination of the processing sites of an Arabidopsis 2S albumin and characterization of the complete gene family.
A:Reference number: JA0161

A:Accession: JA0162
A:Molecule type: DNA
A:Residues: 1-170 <KRE>
A:Cross-references: GB:M22034; NID:g166610; PIDN:AAA32744.1; PID:g166615
R:Conceicao, A.D.S.; Krebers, E.
submitted to the EMBL Data Library, July 1993

A:Description: Tentative title: a cotyledon regulatory region is responsible for the different spatial expression patterns of Arabidopsis 2S albumin genes.
A:Reference number: S34674
A:Accession: S34677
A:Molecule type: DNA

A:Residues: 1-170 <CON>
A:Cross-references: EMBL:Z24745; NID:g395203; PIDN:CAA80871.1; PID:g395205
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voel, M.; Robben, J.; Volckaert, G.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, March 1999
A:Reference number: Z15484
A:Accession: T06045

A:Molecule type: DNA
A:Residues: 1-170 <BEV>
A:Cross-references: EMBL:AL035680; GSPDB:GN00062; ATSP:T24A18.100
A:Experimental source: cultivar Columbia; BAC clone T24A18
C:Genetics:

A:Gene: T24A18.100
A:Map position: 4
A:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: seed; storage protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-170/Product: 2S albumin 2 precursor #status predicted <AR2>
F:38-72/Product: 2S albumin 2 small chain #status predicted <SMC>
F:89-169/Product: 2S albumin 2 large chain #status predicted <LAC>

NMW2 Length: 170 September 12, 2002 12:08 Type: P Check: 3182
Found using 'consensus' (consens.key)

1 MAKLFPLFATLALFLLANASIQTTVIEVDEEDNQLRCCROFLQHLRACQRFTHRR
45

61 LMRWQMGQGGGSLDDEFDLEDIEDNPOGQGHQILQCCSESLQEBPVCVCPFLRQ

121 AARAVSLQGGHGPQSRRIKYTKATLPNICKIQVGCPCPTPIPPPPY
158

1 match found in sequence:
nmw3 : TOIG of: nmw3 check: 419 from: 1 to: 164
(from "consensus_pir.pep")
TOIG of: nmw3 check: 419 from: 1 to: 164

P1:NMW3 - 2S albumin 3 precursor - Arabidopsis thaliana
N:Alternate names: seed storage protein AT253
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 10-May-2001
C:Accession: JA0163; S34674; PNO174; T06046
R:Krebers, E.; Herdies, L.; de Clercq, A.; Seurinck, J.; Leemans, J.; Van Damme, J.; Segura, M.; Gheysen, G.; Van Montagu, M.; Vandekerckhove, J.
Plant Physiol. 87, 859-866, 1988
A:Title: Determination of the processing sites of an Arabidopsis 2S albumin and characterization of the complete gene family.
A:Reference number: JA0161
A:Accession: JA0163

A:Molecule type: DNA
A:Residues: 1-164 <KRE>
A:Cross-references: GB:M22035; NID:g166611; PIDN:AAA32745.1; PID:g166616
R:Conceicao, A.D.S.; Krebers, E.
submitted to the EMBL Data Library, July 1993
A:Description: Tentative title: a cotyledon regulatory region is responsible for the different spatial expression patterns of Arabidopsis 2S albumin genes.

A:Reference number: S34674
A:Accession: S34674
A:Molecule type: DNA
A:Residues: 1-164 <CON>
A:Cross-references: EMBL:Z24744; NID:g395200; PIDN:CAA80868.1; PID:g395201
R:Tangita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.
submitted to JIPID, December 1995

A:Description: Two dimensional electrophoresis of plant proteins and standardization of the gel patterns.
 A:Reference number: PNO173
 A:Accession: PNO174
 A:Molecule type: Protein
 A:Residues: 81-94 <TSU>
 A:Experimental source: seeds
 R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schueller, C. submitted to the Protein Sequence Database, March 1999
 A:Reference number: 215484
 A:Accession: T06046
 A:Molecule type: DNA
 A:Residues: 1-164 <BEV>
 A:Cross-references: EMBL:AL035680; GSPDB:GN00062; ATSP:T24A18.110
 A:Experimental source: cultivar Columbia; BAC clone T24A18
 C:Genetics:
 A:Gene: T24A18.110
 A:Map position: 4
 C:Superfamily: wheat alpha-amylase inhibitor
 C:Keywords: seed; storage protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-164/Product: 2S albumin 3 proprotein #status predicted <AT2>
 F:38-72/Product: 2S albumin 3 small chain #status predicted <SMC>
 F:82-162/Product: 2S albumin 3 large chain #status predicted <LAC>
 MM03 Length: 164 September 12, 2002 12:08 Type: P Check: 419 ..
 Found using 'consensus' (consens.key)

1 MANKLFLVCATLALCFLLTNASIVRTVEFEEDDASNPVPROQCKEFOQSQHLRACOR
 45
 61 WMSKMRGCGGSPSLDPEFDEPGQGTQLQGCCNELRQEPVCVCTLKQARANSLS
 121 QGQHGPFQSRKTYQSAKYLPNICKIQGVEGCFQRTTPEPPYY
 151

-----1-----
 2 matches found in sequence:
 rcs ; TOIG of: rcs check: 1269 from: 1 to: 258
 (from "consensus_plr.pep")
 TOIG of: rcs check: 1269 from: 1 to: 258

P1:RCS - 2S seed storage protein precursor - castor bean
 N:Alternate names: 2S albumin precursor
 C:Species: Ricinus communis (castor bean)
 C>Date: 14-Nov-1983 #sequence-revision 08-Feb-1996 #text-change 18-Jun-1999
 C:Accession: S11499; S11500; S11501; S11502; S27221; A01328; S27222
 R:Irwin, S.D.; Lord, J.M.
 Nucleic Acids Res. 18, 5890, 1990
 A:Title: Nucleotide sequence of a Ricinus communis 2S albumin precursor gene.
 A:Reference number: S11499; MUID:91016940
 A:Accession: S11499
 A:Molecule type: DNA
 A:Residues: 1-258 <IRW>
 A:Cross-references: EMBL:X54158; NID:g21067; PIDN:CAA38097.1; PID:g21068
 A:Note: the authors translated the codon CTC for residue 14 as Phe, CCA for residue 74 as Thr and CAA for residue 194 as Asp
 R:Irwin, S.D.; Keen, J.N.; Findlay, J.B.C.; Lord, J.M.
 Mol. Gen. Genet. 222, 400-408, 1990
 A:Title: The Ricinus communis 2S albumin precursor: a single preproprotein may be processed into two different heterodimeric storage proteins.
 A:Reference number: S11500; MUID:91109729
 A:Accession: S11500
 A:Molecule type: mRNA
 A:Residues: 1-13, 'F', 15-73, 'T', 75-258 <IR2>
 A:Experimental source: clone 14g4
 A:Accession: S11501
 A:Molecule type: mRNA

A:Residues: 'SFAIVTVF', 15-73, 'T', 75-258 <IR3>
 A:Experimental source: clone 8g8
 A:Accession: S11502
 A:Molecule type: mRNA
 A:Residues: 'M', 4, 'LS', 7-13, 'F', 15-21 <IRF>
 A:Experimental source: clone 10a12
 A:Accession: S27221
 A:Molecule type: protein
 A:Residues: 'X', 37, 'X', 39-45, 'X', 158-161, 'X', 163-174, 'X' <IRA>
 R:Sharief, F.S.; Li, S.S.L.
 J. Biol. Chem. 257, 14753-14759, 1982
 A:Title: Amino acid sequence of small and large subunits of seed storage protein from Ricinus communis.
 A:Reference number: A92357; MUID:83082772
 A:Accession: A01328
 A:Molecule type: Protein
 A:Residues: 157-190; 194-221, 'Q', 223-225, 230-233, 'N', 235-254, 'Q', 256-258 <SHA>
 A:Note: 230-Ser was also found
 A:Note: there is considerable similarity between residues 181-231 of this protein and residues 1-47 of the lima bean Bowman-Birk protease inhibitor
 R:Odani, S.; Koide, T.; Ono, T.; Ohnishi, K.
 Biochem. J. 213, 343-345, 1983
 A:Title: Structural relationship between barley (Hordeum vulgare) trypsin inhibitor and castor bean (Ricinus communis) storage protein.
 A:Reference number: A90322; MUID:83308577
 A:Contents: annotation
 A:Note: this protein is homologous with trypsin inhibitor from barley
 C:Complex: consists of two chains linked by two disulfide bonds involving Cys-162 and Cys-175
 C:Superfamily: 2S seed storage protein precursor
 C:Keywords: pyroglutamic acid; seed; storage protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-35/Domain: propeptide #status predicted <PRO>
 F:36-72/Product: probable 2S seed storage protein small chain 2 #status experimental <SCH1>
 F:87-156/Product: probable 2S seed storage protein large 2 #status predicted <LCH1>
 F:157-190/Product: 2S seed storage protein small chain #status experimental <SMU>
 F:194-258/Product: 2S seed storage protein large chain #status experimental <LRG>
 F:194/Modified site: pyrrolidone carboxylic acid (gin) (in mature form) #status experimental
 R:CS Length: 258 September 12, 2002 12:08 Type: P Check: 1269 ..
 Found using 'consensus' (consens.key)

1 MAKLIPTALVSLFIITANSAFAYRTTITTEIDSKGEREGSSQCRQEVQRKDLSS
 49
 61 CERYLRQSSRRSPGEVIRMPGDENQOESQLOCCNQVQVDECCCEAIKIYAEQO
 121 IQQGLHGESESEVNAQRAGEIVSSGVCRCMROTRTPNSQSGRGQIQEQONLRQCEYIK
 149 162
 181 QQVSGQGRRRSDNQRSLRGCCDHLKOMQSQRCRGLRQAIRQOOSQGLQGDVFEAPR
 241 TAAALPMSGVSPTECRF
 256

-----1-----
 2 matches found in sequence:
 S01062 ; TOIG of: S01062 check: 2705 from: 1 to: 295
 (from "consensus_plr.pep")
 TOIG of: S01062 check: 2705 from: 1 to: 295

P1:S01062 - 2S seed storage protein precursor (clone Hag5) - common sunflower
 N:Alternate names: 2S albumin storage protein
 C:Species: Helianthus annuus (common sunflower)
 C>Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Jul-2000
 C:Accession: S01062
 R:Allen, R.D.; Cohen, E.A.; Vonder Haar, R.A.; Adams, C.A.; Ma, D.P.; Nessler, C.L.; Thomas, T.L.
 M.L. Gen. Genet. 210, 211-218, 1987
 A>Title: Sequence and expression of a gene encoding an albumin storage protein in sunflower.
 A:Reference number: S01062; MUID:88142538
 A:Accession: S01062
 A:Molecule type: DNA
 A:Residues: 1-295 <ALT>
 A:Cross-references: EMBL:X06410; NID:q18799; PIDN:CA92699.1; PID:q18800
 A>Note: part of this sequence, including the amino end of the mature protein, was confirmed by protein sequencing
 C:Genetics:
 A:Introns: 192/2
 C:Superfamily: gliadin
 F:1-20/Domain: signal sequence #status predicted <STG>
 F:162-295/Product: 2S seed storage protein #status experimental <MAT>
 S01062 Length: 295 September 12, 2002 12:08 Type: P Check: 2705
 Found using 'consensus' (consens.key)

128 OOLKOKAOLPNCVLQSRCEIGITTTVTESNIDIPFRDRPFGTSGSQCREIEIQRPV
 178
 178

188 GECRFVEBQMOCSPRSTRPYQQRPGQOQOQOQGLQOQCNELOWKRECHCEALQEVAR

 248 RVMROPQOQOQOQRRGQFGGEMETARRVIONLPNCQCLEVOQCTTCTG
 290
 293

 1 match found in sequence:
 s10029 ; TOIG of: s10029 check: 7810 from: 1 to: 168
 (from "consensus_pir.pep")
 TOIG of: s10029 check: 7810 from: 1 to: 168

P1:S10029 - alpha-amylase inhibitor, tetrameric, chain CM3 precursor - wheat
 C:Species: Triticum aestivum (common wheat)
 C>Date: 08-Jun-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
 C:Accession: S10029
 R:Garcia-Maroto, F.; Marana, C.; Mena, M.; Garcia-Olmedo, F.; Carbonero, P.
 Plant Mol. Biol. 14, 845-853, 1990
 A>Title: Cloning of cDNA and chromosomal location of genes encoding the three types of subunits of the wheat tetrameric inhibitor of insect alpha-amylase.
 A:Reference number: S10027; MUID:91346675
 A:Accession: S10029
 A:Molecule type: mRNA
 A:Residues: 1-168 <GAR>
 A:Cross-references: EMBL:X15754; NID:q21712; PIDN:CA935597.1; PID:q21713
 A:Experimental source: cv. Chinese spring
 C:Complex: tetramer
 C:Superfamily: wheat alpha-amylase inhibitor
 C:Keywords: alpha-amylase inhibitor; tetramer
 F:1-25/Domain: signal sequence #status predicted <STG>
 F:26-168/Product: alpha-amylase inhibitor, tetrameric, chain CM3 #status predicted <MAT>

S10029 Length: 168 September 12, 2002 12:08 Type: P Check: 7810
 Found using 'consensus' (consens.key)

1 MACKSSCLLLAVALLSVLAASASGSCVPGVAFRTNLPFCRDYVYQTCGFTTPSK
 29

61 LPENMTASISVPGKPYLAKLYCCQELAEISQCCRCALRFIALPVPSPQVDPDSGNVG

121 ESQLIDPGCPREMQMDFVRLVAPGQCNIATIHNVRCPAVEPLWT
 159

 1 match found in sequence:
 s14946 ; TOIG of: s14946 check: 1431 from: 1 to: 146
 (from "consensus_pir.pep")
 TOIG of: s14946 check: 1431 from: 1 to: 146

P1:S14946 - 2S seed storage protein large chain - Brazil nut
 N:Alternate names: albumin 2S precursor
 C:Species: Bertholletia excelsa (Brazil nut)
 C>Date: 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 21-Jul-2000
 C:Accession: S14946; S14479; S06252; S21640; B25802
 R:Gander, E.S.; Holmstrom, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro, M.; Grossi de Sa, M.F.
 Plant Mol. Biol. 16, 437-448, 1991
 A>Title: Isolation, characterization and expression of a gene coding for a 2S albumin from Bertholletia excelsa (Brazil nut).
 A:Reference number: S14946; MUID:91370890
 A:Accession: S14946
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-146 <GAN>
 A:Cross-references: EMBL:X54490; NID:q17710; PIDN:CA93862.1; PID:q17711
 A>Note: the authors translated the codon CTT for residue 13 as Val and GTC for residue 14 as Leu
 R:Bassener, R.; Schlesier, B.
 submitted to the EMBL Data Library, December 1990
 A:Reference number: S14479
 A:Accession: S14479
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-146 <BAS>
 A:Cross-references: EMBL:X57027; NID:q17714; PID:q17715
 R:Altendach, S.B.; Pearson, K.W.; Leung, F.W.; Sun, S.S.M.
 Plant Mol. Biol. 8, 239-250, 1987
 A>Title: Cloning and sequence analysis of a cDNA encoding a Brazil nut protein exceptionally rich in methionine.
 A:Reference number: S06252
 A:Accession: S06252
 A>Status: not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-146 <ALT>
 A:Experimental source: clone pHS-3
 A>Note: part of this sequence, including the amino end of the large chain, was confirmed by protein sequencing
 R:Bassener, R.; Schlesier, B.
 submitted to the EMBL Data Library, December 1990
 A:Reference number: S21640
 A:Accession: S21640
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-146 <BA2>
 A:Cross-references: EMBL:X57028; NID:q17716; PID:q17717
 R:Ampe, C.; Van Damme, J.; de Castro, L.A.B.; Sampaio, M.J.A.M.; Van Montagu, M.; Vandekerckhove, J.
 Eur. J. Biochem. 159, 597-604, 1986
 A>Title: The amino-acid sequence of the 2S sulphur-rich proteins from seeds of Brazil nut (Bertholletia excelsa H.B.K.).
 A:Reference number: A91173; MUID:87004679

```

A:Accession: B25802
A:Molecule type: protein
A:Residues: 70-90,'E',92-121,'W',123-125,'L',127-142 <AMP>
C:Genetics:
A:Introns: 60/3
C:Superfamily: wheat alpha-amylase inhibitor
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-46/Domain: propeptide #status predicted <PRO>
F:47-69/Product: seed storage protein small chain #status predicted <SCH>
F:70-146/Product: seed storage protein large chain #status experimental <LCH>

S14946 Length: 146 September 12, 2002 12:08 Type: P Check: 1431 ..
Found using 'consensus' (consens.key)

1 MAKISVAAALLVLMALGHATAFRATVTTTVEEBNDECEQEQOMQOMLSHCNRYMRQO
40
-----|-----
61 MEESPQYTMPRGRMEPHMSECEQLEGWDESCRCEGLRMAMRMQOEMQPRGQMRRMM
-----|-----
121 RLAEINIPSRCNLSPPRCPMGSIAGF
137
-----|-----
1 match found in sequence:
S14947 ; TOIG of: S14947 check: 5446 from: 1 to: 154
(from "consensus-plr.pep")
TOIG of: S14947 check: 5446 from: 1 to: 154

P1:S14947 - 2S albumin - Brazil nut
C:Species: Bertholletia excelsa (Brazil nut)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C:Accession: S14947
R:Gander, E.S.; Holmstroem, K.O.; de Paiva, G.R.; de Castro, L.A.B.; Carneiro,
M.; Grossi de Sa, M.F.
Plant Mol. Biol. 16, 437-448, 1991
A:Title: Isolation, characterization and expression of a gene coding for a 2S
albumin from Bertholletia excelsa (Brazil nut).
A:Reference number: S14946; MUID:91370890
A:Accession: S14947
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-154 <GAN>
A:Cross-references: EMBL:X54491; NID:917712; PIDN:CAA38363.1; PID:g17713
C:Genetics:
A:Introns: 67/3
C:Superfamily: wheat alpha-amylase inhibitor

S14947 Length: 154 September 12, 2002 12:08 Type: P Check: 5446 ..
Found using 'consensus' (consens.key)

1 MAKMSVAAALLLVLGQATAFRTVTTTLEBQENPRGRSQOCCREQMERQOQLNHC
47
-----|-----
61 RMYLRQOMESPYQNPRLRGRGEPHLDECCQLERMDMCRCEGLRMMLRRQREMELO
-----|-----
121 GEOMQRIKRAENLISRCNLSPORCPMGYTAWL
145
-----|-----
1 match found in sequence:
S16031 ; TOIG of: S16031 check: 7810 from: 1 to: 168
(from "consensus-plr.pep")
TOIG of: S16031 check: 7810 from: 1 to: 168

P1:S16031 - alpha-amylase inhibitor, tetrameric, chain CM3 precursor - durum

```

```

wheat
C:Species: Triticum durum (durum wheat)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Jul-1999
C:Accession: S16031
R:Gautier, M.F.; Alary, R.; Lullien, V.; Joudrier, P.
submitted to the EMBL Data Library, June 1991
A:Reference number: S16031
A:Accession: S16031
A:Molecule type: mRNA
A:Residues: 1-168 <GAU>
A:Cross-references: EMBL:X61032; NID:921921; PIDN:CAA43367.1; PID:g21922
A:Experimental source: cultivar Agathe; tissue type seed
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: alpha-amylase inhibitor; tetramer
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-168/Product: alpha-amylase inhibitor, tetrameric, chain CM3 #status
predicted <MAT>

S16031 Length: 168 September 12, 2002 12:08 Type: P Check: 7810 ..
Found using 'consensus' (consens.key)

1 MACKSSCSLLILAAVLSVLAASASGSCVPGVAFRTNLLPHCRDYLVLQGTGFTPEGK
29
-----|-----
61 LPEWMTSASITSPCKPYIAKLYCCQELAEISQRCCEALRFTALPVPSPQVDRSGNWG
-----|-----
121 ESGHIDLPGRQMDPVRLLVAPGQCNLATIHNVRCYCPAVEQPLWT
159
-----|-----
1 match found in sequence:
S18241 ; TOIG of: S18241 check: 6711 from: 1 to: 143
(from "consensus-plr.pep")
TOIG of: S18241 check: 6711 from: 1 to: 143

P1:S18241 - alpha-amylase inhibitor, tetrameric, chain CM17 precursor - wheat
C:Species: Triticum aestivum (common wheat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S18241
R:Lullien, V.; Alary, R.; Guirao, A.; Joudrier, P.; Gautier, M.F.
Plant Mol. Biol. 17, 1081-1082, 1991
A:Title: Isolation and nucleotide sequence of a cDNA clone encoding the bread
wheat (Triticum aestivum L.) CM17 protein.
A:Reference number: S18241; MUID:92032767
A:Accession: S18241
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-143 <LUD>
A:Cross-references: EMBL:X59791; NID:921710; PIDN:CAA42453.1; PID:g21711
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: alpha-amylase inhibitor; tetramer

S18241 Length: 143 September 12, 2002 12:08 Type: P Check: 6711 ..
Found using 'consensus' (consens.key)

1 MASKSNYNLFTALLVFIFAAVAAGNDCTPWTSTLITPLPSCRNVVEEQACRIEMGP
30
-----|-----
61 PYLAKQCCQELANIPQOCRCQALRYFMGPKSRDQGLMELPCCPREVMQNFVPIITVP
-----|-----
121 GYCNLTTVHNTPYCLGMBESQWS
134
-----|-----
1 match found in sequence:

```

S20350 : TOIG of: s20350 check: 5085 from: 1 to: 110
 (from "consensus_pir.pep")
 TOIG of: s20350 check: 5085 from: 1 to: 110

P1:S20350 - napin nla - rape
 C:Species: Brassica napus (rape)
 C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 04-Oct-1996
 C/Accession: S20350; S20351
 R:Monsalve, R.I.; Lopez-Otin, C.; Villalba, M.; Rodriguez, R.
 FEBS Lett. 295, 207-210, 1991
 A:Title: A new distinct group of 2 S albumins from rapeseed. Amino acid
 sequence of two low molecular weight napins.
 A:Reference number: S20350; MUID:92111741
 A:Accession: S20350
 A:Molecule type: protein
 A:Residues: 1-110 <MON>
 C:Superfamily: wheat alpha-amylase inhibitor
 C:Keywords: pyroglutamic acid; seed; storage protein
 F:1-31/Product: napin small chain #status experimental <SMA>
 F:32-110/Product: napin large chain #status experimental <LAR>
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status
 experimental

S20350 Length: 110 September 12, 2002 12:08 Type: P Check: 5085 ..
 Found using 'consensus' (consens.key)

1
 5
 QPKCGEFGQEHRLRACQGMIRQLAGSPFGSGPGLRECCNELVQEDQVCVCP

-----|-----
 61 LKQAKSVRVGGHGFQSTRIVQIAKNLPMVCMKQIGTCPTAIPFPF 101

1 match found in sequence:
 S26636 : TOIG of: s26636 check: 79 from: 1 to: 106
 (from "consensus_pir.pep")
 TOIG of: s26636 check: 79 from: 1 to: 106

P1:S26636 - napin nrb - rape
 C:Species: Brassica napus (rape)
 C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 04-Oct-1996
 C/Accession: S26636
 R:Monsalve, R.I.; Lopez-Otin, C.; Villalba, M.; Rodriguez, R.
 FEBS Lett. 295, 207-210, 1991
 A:Title: A new distinct group of 2 S albumins from rapeseed. Amino acid
 sequence of two low molecular weight napins.
 A:Reference number: S20350; MUID:92111741
 A:Accession: S26636
 A:Molecule type: protein
 A:Residues: 1-106 <MON>
 C:Superfamily: wheat alpha-amylase inhibitor
 C:Keywords: pyroglutamic acid; seed; storage protein
 F:1-31/Product: napin small chain #status experimental <SMA>
 F:32-106/Product: napin large chain #status experimental <LAR>
 F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status
 experimental

S26636 Length: 106 September 12, 2002 12:08 Type: P Check: 79 ..
 Found using 'consensus' (consens.key)

1
 5
 QPKCGEFGQEHRLRACQGMIRQLAGSPFGSGPGLRECCNELVQEDQVCVCP

-----|-----
 61 LKQAKSVRVGGHGFQSTRIVQIAKNLPMVCMKQIGTCPTAIP 101

1 match found in sequence:
 S29002 : TOIG of: s29002 check: 4419 from: 1 to: 121

(from "consensus_pir.pep")
 TOIG of: s29002 check: 4419 from: 1 to: 121

P1:S29002 - trypsin inhibitor - rye
 C:Species: Secale cereale (rye)
 C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
 C/Accession: S29002
 R:Lyons, A.; Richardson, M.; Tatham, A.S.; Shewry, P.R.
 Biochim. Biophys. Acta 915, 305-313, 1987
 A:Title: Characterization of homologous inhibitors of trypsin and alpha-amylase
 from seeds of rye (Secale cereale L.).
 A:Reference number: S29002
 A:Accession: S29002
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-121 <LYO>
 C:Superfamily: wheat alpha-amylase inhibitor

S29002 Length: 121 September 12, 2002 12:08 Type: P Check: 4419 ..
 Found using 'consensus' (consens.key)

1
 6
 SVGGQVPLGAMPNPLGACRTYVSOICHVGRPLFTWDMKRRCCDELLAIPAYRCEAL

-----|-----
 61 RILMDGVVTGGVFEESGYLKDMPNCPRVTRGSAATLVAPQECNLPTIHGSPYCPFLQAG 114

121 Y

1 match found in sequence:
 S46514 : TOIG of: s46514 check: 8339 from: 1 to: 148
 (from "consensus_pir.pep")
 TOIG of: s46514 check: 8339 from: 1 to: 148

P1:S46514 - puroindoline-b precursor - wheat
 C:Species: Triticum aestivum (common wheat)
 C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 20-Aug-1999
 C/Accession: S46514
 R:Gautier, M.F.; Aleman, M.E.; Gutrao, A.; Marion, D.; Joudrier, P.
 Plant Mol. Biol. 25, 43-57, 1994
 A:Title: Triticum aestivum puroindolines, two basic cysteine-rich seed proteins:
 cDNA sequence analysis and developmental gene expression.
 A:Reference number: S46514; MUID:94272013
 A:Accession: S46514
 A:Molecule type: mRNA
 A:Residues: 1-148 <GAU>
 A:Cross-references: EMBL:X69912; NID:9509087; PIDN:CAA49537.1; PID:9509088
 C:Superfamily: glycinnin
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-145/Product: puroindoline-a #status predicted <MAT>
 F:146-148/Domain: carboxyl-terminal propeptide #status predicted <PRO>

S46514 Length: 148 September 12, 2002 12:08 Type: P Check: 8339 ..
 Found using 'consensus' (consens.key)

1
 40
 MKTLFLALLALVASTTFAQYSEVGGYNEVGGGGGSGQCPOGRPALSSCKDYMERCF

-----|-----
 61 MKFPPTWPTKMMKGGGEHVRKCKQLSQLAPQRCDSIRRVIGRLGGLGIRGEV 140

121 FKOLRAOSLPSKCMGADCKFPSSGYW 140

1 match found in sequence:
 S48176 : TOIG of: s48176 check: 3743 from: 1 to: 104

(from "consensus_pir.pep")
TOIG of: S48176 check: 3743 from: 1 to: 104
F1:S48176 - mablinin I-1 - Yunnan caper (fragments)
C:Species: Capparis masakal (Yunnan caper)
C:Date: 15-Jul-1995 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998
R:Nirasawa, S.; Nishino, T.; Katahira, M.; Uesugi, S.; Hu, Z.; Kurihara, Y.
Eur. J. Biochem. 223, 989-995, 1994
A:Title: Structures of heat-stable and unstable homologues of the sweet protein mablinin. The difference in the heat stability is due to replacement of a single amino acid residue.
A:Reference number: S48176; MUID:94333405
A:Accession: S48176
A:Molecule type: protein
A:Residues: 1-32,33-104 <NIR>
A:Experimental source: seed
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: albumin; storage protein; sweet taste
F:1-32/Product: mablinin I-1 chain A (fragment) #status experimental <CHA>
F:33-104/Product: mablinin I-1 chain B #status experimental <CHB>
F:4-53,17-42,43-91,55-99/Disulfide bonds: #status experimental
S48176 Length: 104 September 12, 2002 12:08 Type: P Check: 3743 ..
Found using 'consensus' (consens.key)

1 EPLCRPOQHOLRACQRYLRRAQRGIVDEQGPALRLCCNQLRVNKPVCVPLRQ
4
61 AAHQIYOGQIEGPRQVRLFRARNLPNICKIPAVGRQCFTRW
99

1 match found in sequence:
S48178 ; TOIG of: S48178 check: 3177 from: 1 to: 104
(from "consensus_pir.pep")
TOIG of: S48178 check: 3177 from: 1 to: 104
F1:S48178 - mablinin III - Yunnan caper (fragments)
C:Species: Capparis masakal (Yunnan caper)
C:Date: 15-Jul-1995 #sequence_revision 13-Mar-1998 #text_change 20-Mar-1998
R:Nirasawa, S.; Nishino, T.; Katahira, M.; Uesugi, S.; Hu, Z.; Kurihara, Y.
Eur. J. Biochem. 223, 989-995, 1994
A:Title: Structures of heat-stable and unstable homologues of the sweet protein mablinin. The difference in the heat stability is due to replacement of a single amino acid residue.
A:Reference number: S48176; MUID:94333405
A:Accession: S48178
A:Molecule type: protein
A:Residues: 1-32,33-104 <NIR>
A:Experimental source: seed
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: albumin; storage protein; sweet taste
F:1-32/Product: mablinin III chain A (fragment) #status experimental <CHA>
F:33-104/Product: mablinin III chain B #status experimental <CHB>
F:4-53,17-42,43-91,55-99/Disulfide bonds: #status predicted
S48178 Length: 104 September 12, 2002 12:08 Type: P Check: 3177 ..
Found using 'consensus' (consens.key)

1 match found in sequence:
S48180 ; TOIG of: S48180 check: 9891 from: 1 to: 100
(from "consensus_pir.pep")
TOIG of: S48180 check: 9891 from: 1 to: 100
F1:S48180 - mablinin IV - Yunnan caper (fragments)
C:Species: Capparis masakal (Yunnan caper)
C:Date: 15-Jul-1995 #sequence_revision 13-Mar-1998 #text_change 20-Mar-1998
R:Nirasawa, S.; Nishino, T.; Katahira, M.; Uesugi, S.; Hu, Z.; Kurihara, Y.
Eur. J. Biochem. 223, 989-995, 1994
A:Title: Structures of heat-stable and unstable homologues of the sweet protein mablinin. The difference in the heat stability is due to replacement of a single amino acid residue.
A:Reference number: S48176; MUID:94333405
A:Accession: S48180
A:Molecule type: protein
A:Residues: 1-28,29-100 <NIR>
A:Experimental source: seed
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: albumin; storage protein; sweet taste
F:1-28/Product: mablinin IV chain A (fragment) #status experimental <CHA>
F:29-100/Product: mablinin IV chain B #status experimental <CHB>
F:4-49,17-38,39-87,51-95/Disulfide bonds: #status predicted
S48180 Length: 100 September 12, 2002 12:08 Type: P Check: 9891 ..
Found using 'consensus' (consens.key)

1 EPLCRPOQHOLRACQRYLRRAQRGIVDEQGPALRLCCNQLRVNKPVCVPLRQAHQ
4
61 QLYOGIEGPRQVRLFRARNLPNICKIPAVGRQCFTRW
95

1 match found in sequence:
S65447 ; TOIG of: S65447 check: 2875 from: 1 to: 127
(from "consensus_pir.pep")
TOIG of: S65447 check: 2875 from: 1 to: 127
P1:S65447 - allergen Sin a I - white mustard
N:Alternate names: allergen Sin a I large chain
C:Species: Sinapis alba (white mustard)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
R:Gonzalez de la Pena, M.A.; Monsalve, R.I.; Batanero, E.; Villalba, M.; Rodriguez, R.
Eur. J. Biochem. 237, 827-832, 1996
A:Title: Expression in Escherichia coli of Sin a I, the major allergen from mustard.
A:Reference number: S65447; MUID:96235251
A:Accession: S65447
A:Molecule type: protein
A:Residues: 1-127 <GON>
A:Experimental source: seed
R:Mendez-Arias, L.; Moneo, I.; Dominguez, J.; Rodriguez, R.
Eur. J. Biochem. 177, 159-166, 1988
A:Title: Primary structure of the major allergen of yellow mustard (Sinapis alba L.) seed, sin a I
A:Reference number: S01791; MUID:89030681
A:Accession: S01792
A:Molecule type: protein
A:Residues: 40-127 <MEN>
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: seed; storage protein
S65447 Length: 127 September 12, 2002 12:08 Type: P Check: 2875 ..
Found using 'consensus' (consens.key)

```

1 -----
PAGEFRIPKREKREQOAHOLRACQOQWHLKQAMQSGSPQGPQQRPLLQCCNELHQE
10
61 -----
EPLVCPTLKGASKAVKQVROQLEQGGQGGPHVISRIYQTATHLPRVCNIPQVSVCFK
117
121 KTMGPGS
-----
1 match found in sequence:
S78525 : TOIG of: S78525 check: 8003 from: 1 to: 171
(from "consensus_pir.pep")
TOIG of: S78525 check: 8003 from: 1 to: 171

P1:S78525 - alpha-amylase inhibitor, tetrameric, chain CMD precursor - barley
N:Alternate names: seed protein CMD
C:Species: Hordeum vulgare (barley)
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 16-Jul-1999
C:Accession: S78525; JAO071
R:Medina, J.; Hueros, G.; Carbonero, P.
Plant Mol. Biol. 23, 535-542, 1993
A:Title: Cloning of cDNA, expression, and chromosomal location of genes
encoding the three types of subunits of the barley tetrameric inhibitor of
insect alpha-amylase.
A:Reference number: S39560; MUID:94033333
A:Accession: S78525
A:Molecule type: mRNA
A:Residues: 1-171 <MED>
A:Cross-references: EMBL:X69939; NID:9452324; PIDN:CAA9557.1; PID:9452325
R:Halford, N.G.; Morris, N.A.; Urwin, P.; Williamson, M.S.; Kasarda, D.D.; Lew,
E.J.L.; Kreis, M.; Shewry, P.R.
Biochim. Biophys. Acta 950, 435-440, 1988
A:Title: Molecular cloning of the barley seed protein CMD: a variant member of
the alpha-amylase/trypsin inhibitor family of cereals.
A:Reference number: JAO071; MUID:89000796
A:Accession: JAO071
A:Molecule type: mRNA
A:Residues: 12-171 <HAL>
A:Cross-references: EMBL:X13198; NID:921716; PIDN:CAA31585.1; PID:9758343
A:Note: parts of this sequence, including the amino end of the mature protein,
were confirmed by protein sequencing
C:Genetics:
A:Gene: iat3
A:Map position: 4H
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: alpha-amylase inhibitor; seed; tetramer
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-171/Product: alpha-amylase inhibitor, tetrameric, chain CMD #status
experimental <MAT>

S78525 Length: 171 September 12, 2002 12:08 Type: P Check: 8003
Found using 'consensus' (consens.key)

1 -----
MACKSRKILLLATVWVSFAAAAAAATDCSPGVAFFTNLGHCRDYLQOTCAVETP
32
61 -----
GSKLPEWMTSAELNYPQPYLAKLYCCQELAEIPQCCREALRYFALPVPSPQVDPSTG
-----
121 NVQSGSLMDLPQCPREKQRFVRLVAPGQCNLATIHNVRYCAVDEPLMI
162
-----
1 match found in sequence:
T04470 : TOIG of: T04470 check: 7816 from: 1 to: 147
(from "consensus_pir.pep")
TOIG of: T04470 check: 7816 from: 1 to: 147

```

```

P1:T04470 - probable trypsin inhibitor Cme2.2 - barley
C:Species: Hordeum vulgare (barley)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C:Accession: T04470
R:Rojo, J.
submitted to the EMBL Data Library, January 1998
A:Reference number: Z15362
A:Accession: T04470
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-147 <ROY>
A:Cross-references: EMBL:AJ223458; PIDN:CAA11373.1
C:Genetics:
A:Gene: It1
C:Superfamily: wheat alpha-amylase inhibitor,
C:Keywords: serine proteinase inhibitor

T04470 Length: 147 September 12, 2002 12:08 Type: P Check: 7816
Found using 'consensus' (consens.key)

1 -----
MAFKYQLLSAAVMATLATVTSFGDMCAPGDALPANPLRACRTYVVSQICHVGRLLST
29
61 -----
WDMKRRCDELSAIPAYCRCEALRIIMDGTVQGVGAYFKDMPNSPRVQTSYANLV
-----
121 NPOECNIMTIHGSPSCPELOPGYEVVL
136
-----
1 match found in sequence:
T09850 : TOIG of: T09850 check: 6563 from: 1 to: 139
(from "consensus_pir.pep")
TOIG of: T09850 check: 6563 from: 1 to: 139

P1:T09850 - albumin 2S storage protein precursor - upland cotton
C:Species: Gossypium hirsutum (upland cotton)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C:Accession: T09850
R:Galau, G.A.; Wang, H.Y.C.; Hughes, D.W.
submitted to the EMBL Data Library, January 1992
A:Description: cotton Mat5 (C164) gene and cDNAs encoding a methionine-rich 2S
albumin storage protein.
A:Reference number: Z16886
A:Accession: T09850
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-139 <GAL>
A:Cross-references: EMBL:M83301; NID:9167310; PID:9167311
C:Genetics:
A:Gene: Mat5-D
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-139/Product: albumin 2S storage protein #status predicted <MAT>

T09850 Length: 139 September 12, 2002 12:08 Type: P Check: 6563
Found using 'consensus' (consens.key)

1 -----
MAKLAVLATLALILFLANASITSVTESEENRDSCEQDIRKQAHKHCQKYEEWGE
36
61 -----
GSDNIAGYIDSCQQLKKMDTQCRQGLRHATWQOMQOMQOMGSKOMREIMQKTKNI
-----
121 MSECMEPEKCDPRSRSLI
131
-----

```


1 match found in sequence:
 T09878 ; TOIG of: t09878 check: 6011 from: 1 to: 139
 (from "consensus_pir.pep")
 TOIG of: t09878 check: 6011 from: 1 to: 139

P1:T09878 - albumin 2S storage protein precursor Mat5-A - upland cotton
 C:Species: Gossypium hirsutum (upland cotton)
 C>Date: 16-Jul-1999 #sequence.revision 16-Jul-1999 #text_change 29-Oct-1999
 C:Accession: T09878
 R:Galau, G.A.; Wang, H.Y.C.; Hughes, D.W.
 Submitted to the EMBL Data Library, January 1992
 A:Description: Cotton Mat5-A (C164) gene and Mat5-D cDNAs encoding
 methionine-rich 2S albumin storage proteins.
 A:Reference number: 216893
 A:Accession: T09878
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-139 <GAL>
 A:Cross-references: EMBL:M86213; NID:g167358; PID:g167359
 C:Genetics:
 A:Gene: Mat5-A
 C:Keywords: storage protein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-139/Product: albumin 2S storage protein Mat5-A #status predicted <MAT>
 T09878 length: 139 September 12, 2002 12:08 Type: P Check: 6011 ..
 Found using 'consensus' (consens.key)

1 MAKLAVALATLILFLANASITSVTWESEBNDSCSQIIRKQALHKQCKYMEELGGE
 36

61 GSDNIAGYIDSCQQLKMDTQCRQGLRHATMQOMQOMQOMGSKOMREIMQVTKKI

121 -----|-----
 MSECEMERGCDTPSRSLI
 131

1 match found in sequence:
 t1bh ; TOIG of: t1bh check: 7837 from: 1 to: 148
 (from "consensus_pir.pep")
 TOIG of: t1bh check: 7837 from: 1 to: 148

P1:TI1BH - trypsin inhibitor CME precursor - barley
 C:Species: Hordeum vulgare (barley)
 C>Date: 14-Nov-1983 #sequence.revision 12-Apr-1996 #text_change 16-Jun-2000
 C:Accession: S21451; A01325; J00342; S71590
 R:Royo, J.
 Submitted to the EMBL Data Library, April 1992
 A:Reference number: S21451
 A:Accession: S21451
 A:Molecule type: DNA
 A:Residues: 1-148 <ROY>
 A:Cross-references: EMBL:X65875; NID:g19008; PIDN:CAA46705.1; PID:g19009
 R:Odani, S.; Koide, T.; Ono, T. 1983
 J. Biol. Chem. 258, 7998-8003, 1983
 A:Title: The complete amino acid sequence of barley trypsin inhibitor.
 A:Reference number: A01325; MUID:83238397
 A:Accession: A01325
 A:Molecule type: protein
 A:Residues: 25-145 <ODA>
 A>Note: residue 57 is tentatively assigned as the site of interaction with
 trypsin
 R:Rodriguez-Palenzuela, P.; Royo, J.; Gomez, L.; Sanchez-Monge, R.; Salcedo,
 G.; Molina-Cano, J.L.; Garcia-Olmedo, F.; Carbonero, P.
 Mol. Gen. Genet. 219, 474-479, 1989
 A:Title: The gene for trypsin inhibitor CME is regulated in trans by the lys 3a
 locus in the endosperm of barley (Hordeum vulgare L.).
 A:Reference number: J00342; MUID:90158508
 A:Accession: J00342

A:Molecule type: mRNA
 A:Residues: 1-142, 'A', 144 <ROD>
 A:Cross-references: GB:X17302; NID:g18956; PIDN:CAA35188.1; PID:g1405736
 A>Note: the authors translated the codon GAG for residue 111 as Asp and GCA for
 residue 143 as Gly
 R:Royo, J.; Diaz, I.; Rodriguez-Palenzuela, P.; Carbonero, P.
 Plant Mol. Biol. 31, 1051-1059, 1996
 A:Title: Isolation and promoter characterization of barley gene tr1 encoding
 trypsin inhibitor Brl-CME: differential activity in wild-type and mutant lys3a
 endosperm.
 A:Reference number: S71590; MUID:97000917
 A:Accession: S71590
 A:Molecule type: DNA
 A:Residues: 1-148 <RO2>
 A:Cross-references: EMBL:X65875; NID:g19008; PIDN:CAA46705.1; PID:g19009
 C:Genetics:
 A:Gene: tr1
 C:Superfamily: wheat alpha-amylase inhibitor
 C:Keywords: serine proteinase inhibitor
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-144/Product: trypsin inhibitor CME #status predicted <MAT>
 F:57/Inhibitory site: Arg (trypsin) #status predicted
 TI1BH length: 148 September 12, 2002 12:08 Type: P Check: 7837 ..
 Found using 'consensus' (consens.key)

1 MAFRYOLLNAAVAILVATATSFSGDCAPGDALPHNPLRACRYVVSQICHQPRLLT
 29

61 SDMKRRCCDELSAIPAYCRCALRIIMGVYTWGAFEGAFKSPNPRRQRTSYANL

121 -----|-----
 VTPQECNIGTIHGSAYCPQLPGYGVVL
 137

1 match found in sequence:
 t1zm1 ; TOIG of: t1zm1 check: 5547 from: 1 to: 155
 (from "consensus_pir.pep")
 TOIG of: t1zm1 check: 5547 from: 1 to: 155

P1:TI1ZM1 - trypsin/factor X11a inhibitor precursor - maize
 N:Alternate names: Hageman factor inhibitor
 C:Species: Zea mays (maize)
 C>Date: 28-Aug-1985 #sequence.revision 31-Dec-1992 #text_change 18-Jun-1999
 C:Accession: S20850; A01327
 R:Men, L.; Huang, J.K.; Zou, K.C.; Johnson, B.H.; Muthukrishnan, S.; Mackay,
 V.; Manney, T.R.; Manney, M.; Reek, G.R.
 Plant Mol. Biol. 18, 813-814, 1992
 A:Title: Nucleotide sequence of a cDNA clone that encodes the maize inhibitor
 of trypsin and activated Hageman factor.
 A:Reference number: S20850; MUID:92216060
 A:Accession: S20850
 A:Molecule type: mRNA
 A:Residues: 1-155 <MEN>
 A:Cross-references: EMBL:X54064; NID:g22326; PIDN:CAA37998.1; PID:g22327
 R:Mahoney, W.C.; Herndonson, M.A.; Jones, B.; Powers, D.D.; Corfman, R.S.;
 Reek, G.R.
 J. Biol. Chem. 259, 8412-8416, 1984
 A:Title: Amino acid sequence and secondary structural analysis of the corn
 inhibitor of trypsin and activated Hageman factor.
 A:Reference number: A01327; MUID:84239823
 A:Accession: A01327
 A:Molecule type: protein
 A:Residues:
 29-48, 'C', 50-55, 'R', 57-60, 'PR', 61-106, 'A', 108-116, 'A', 118-119, 'Q', 121-134, 'E', 13
 6-138 <MAT>
 A>Note: 117-Glu was also found
 C:Comment: Five disulfide bonds are present.
 C:Superfamily: wheat alpha-amylase inhibitor

C;Keywords: disulfide bond; serine proteinase inhibitor
F;1-28/Domain: signal sequence #status predicted <SIG>
F;29-138/Product: trypsin/factor Xlla inhibitor #status experimental <MAT>
F;139-155/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;62/Inhibitory site: Arg (trypsin) #status experimental

TI:MI Length: 155 September 12, 2002 12:08 Type: P Check: 5547 ..
Found using 'consensus' (consens.key)

1 MASSSSSHRRLLIAAVLTVLAASAGTSCVPGMALPHNPLPSCRWYVTSRTCGIG
34

61 PRLPWLKRRCCRELADIPAYCRCTALSTIMDAIPPPDAQLERLEDLPGCPREVQR

121 GFATLVTEACNLATISGVACPMILGGTMPK
143

1 match found in sequence:
wmr219 : TOIG of: wmr219 check: 1252 from: 1 to: 186
(from 'consensus-pir.pep')
TOIG of: wmr219 check: 1252 from: 1 to: 186

P1:WMR219 - 19K globulin precursor - rice
N;Alternate names: alpha-globulin
C;Species: Oryza sativa (rice)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 18-Jun-1999
C;Accession: S20024; S25735; P00497
R;Shorrosh, B.S.; Wen, L.; Zen, K.C.; Huang, J.K.; Pan, J.S.; Hermodson, M.A.;
Tanaka, K.; Muthukrishnan, S.; Reeck, G.R.
Plant Mol. Biol. 18, 151-154, 1992
A;Title: A novel cereal storage protein: molecular genetics of the 19 kDa
globulin of rice.

A;Reference number: S20024; MUID:92119226
A;Accession: S20024
A;Molecule type: mRNA
A;Residues: 1-186 <SHO>
A;Cross-references: EMBL:X63990; NID:g20158; PIDN:CAA5400.1; PID:g20159
A;Accession: S25735
A;Molecule type: protein
A;Residues: 66-74;108-133;171-186 <SH2>
R;Krishnan, H.B.; Pueppke, S.G.
Biochem. Biophys. Res. Commun. 193, 460-466, 1993
A;Title: Nucleotide sequence of an abundant rice seed globulin: homology with
the high molecular weight glutelins of wheat, rye and triticale.

A;Reference number: P00497; MUID:93277591
A;Accession: P00497
A;Molecule type: mRNA
A;Residues: 6-186 <KRI>
A;Cross-references: GB:I12252
A;Experimental source: seed
C;Superfamily: wheat alpha-amylase inhibitor
C;Keywords: storage protein
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-186/Product: 19K globulin #status predicted <MAT>

WMR219 Length: 186 September 12, 2002 12:08 Type: P Check: 1252 ..
Found using 'consensus' (consens.key)

1 MASKVVFFAALMAAVAIISGAHVSEEMFRDRCOREVQDSPLDACRQVLDROLTGRE
36

61 RFQPMFRFGALGLRMQCCOQLDVSRERCALARVWRSYSESMPLLEQGWSSSSEY

121 YGEGSSSEQGYGEGSSSEGYGEGQOQPGMTVRLLTRARQYAAOLPSMCRVEPOQCSI

181 FAAGQY

-- Search Statistics --

Times:	CPU	Total Elapsed
	00:00:00.00	00:00:00.00
Number of sequences searched:	27	
Number of sequence hits:	27	
Number of separate matches:	29	
Number of sequence hits saved:	0	

> 0 <
01/10 Intelligenetics
> 0 <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "consensus-sp" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.

Selected sequence key from "consens.key":

consensus (AA) ID consensus AA preliminary pattern
followed by

1
2 c
2 any character from 9 to 13 times
2 c
2 q or e or r or k
2 any character
2 a or f or g or i or l or m or p or t or v or w or y
2 l or i or v or m
2 any character from 15 to 39 times
2 cc
2 c or d or e or h or k or n or o or q or r or s
2 q or e or h
2 l or v
2 any character
2 any character
2 any character
2 any character
2 any character
2 c
2 any character
2 c
2 any character
2 any character
2 l or i
2 any character from 13 to 56 times
2 g
2 any character from 15 to 26 times
2 c
2 any character
2 v or i or l or m
2 any character from 1 to 8 times
2 c

Selected files:

File : consensus_sp.pep

-- Output Parameters --

Format Options: File Options:
Nucleic acid code matching Exact Indirect file
Find non-matching hits only No Sequence or key file
Report key used Yes List of hits
Note position of hit Yes Hit display
Display full annotations Yes Name and annotations
Sequence context 50 Yes

-- Run Parameters --

Run mode Batch
Time to start comparison now
Notify at end of run No

1 match found in sequence:
Zs1capma ; Mabinlin I-1, A and B chains (MAB I) (Sweet protein).
(from "consensus_sp.pep")
TOIG of: Zs1_capma check: 3743 from: 1 to: 104

ID 2SS1_CAPMA STANDARD: PRT: 104 AA.
AC P80351;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mabinlin I-1, A and B chains (MAB I) (Sweet protein).
OS Capparis masatakai (Mabinling).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Capparis.
OX NCBI_TaxID=13395;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=94333405; PubMed=8055976;
RA Nirasawa S., Nishino T., Katsuhira M., Uesugi S., Hu Z., Kurihara Y.;
RT "Structures of heat-stable and unstable homologues of the sweet
protein mabinlin. The difference in the heat stability is due to
replacement of a single amino acid residue.";
RL Eur. J. Biochem. 223:989-995(1994).
CC -I- FUNCTION: 2S SEED STORAGE PROTEIN HAVING SWEETNESS-INDUCING
CC ACTIVITY. THIS FORM IS NOT HEAT STABLE.
CC -I- SUBUNIT: HETERODIMER OF A SMALL A AND A LARGE B CHAIN LINKED BY
CC DISULFIDE BONDS.
CC -I- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000617; Napin.
DR Pfam: PF01631; Seedstore_2S; 1.
DR PRINTS: PR00496; NAPIN.
DR PRODOM: PD002498; Napin; 1.
DR SMART: SM00499; AAI; 1.
KW Seed storage protein; Albumin; Sweet-taste.
FT CHAIN 1 32 MABINLIN I, A CHAIN.
FT NON_CONS 32 33
FT CHAIN 33 104 MABINLIN I, B CHAIN.
FT DISULFID 4 53
FT DISULFID 17 42
FT DISULFID 43 91
FT DISULFID 55 99
SQ SEQUENCE 104 AA; 12284 MW; 156444AB50D5EE60 CRC64;
2SS1_CAPMA Length: 104 September 12, 2002 12:09 Type: P Check: 3743 ..
Found using 'consensus' (consens.key)
1
4
EPLCRFQOHQHLNACQRYIRRAQRGLVDEQPALRLCCNLRQVKNPCVCPLRQ
61 AAHQOLYOGGLEGRQVROLFRANRLPNICKIPAVGRCQFTRW 99
1 match found in sequence:
2ss2arath: 2S seed storage protein 2 precursor (2S albumin storage protein)
(from "consensus_sp.pep")
TOIG of: 2ss2_arath check: 3182 from: 1 to: 170
ID 2SS2_ARATH STANDARD: PRT: 170 AA.
AC P15458;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 2S seed storage protein 2 precursor (2S albumin storage protein)
DE (NMWU2-2S albumin 2).
GN AT2S2 OR AT4G27150 OR T24A18.100.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA Krebbers E., Herdies L., de Clercq A., Seurinck J., Leemans J.,
 RA van Damme J., Segura M., Gheysen G., van Montagu M.,
 RA Vandekerckhove J.;
 RT "Determination of the processing sites of an Arabidopsis 2S albumin
 RT and characterization of the complete gene family.";
 RL Plant Physiol. 87:859-866(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA Conceicao A.D.S., Krebbers E.;
 RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terryn N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Maché R., Mueller M.,
 RA Kreis M., Delsen M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohnselt J., Zimmermann W., Weiler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weltzenegger T., Bothe G., Ransperger U., Hilbert H., Braun M.,
 RA Holzner E., Brandt A., Peters S., van Staveren N., Dirkse W.,
 RA Woolman P., Klein Lankhorst R., Rose M., Haut J., Koelter P.,
 RA Berneser S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buyschaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Leonard N., McIay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann T.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehmert S.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gebel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzi A.,
 RA Neumann S., Argirion A., Vitale D., Lidgort R., Piravandi E.,
 RA Massenot O., Oudgley F., Clapaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chefor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijman L., Schwarz S., Scholler P., Heber S., Francis P., Bietke C.,
 RA Frisman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threiden J.,
 RA Stoneking T., Kallick J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mxnp P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Perin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghof A., Jones K., Dione K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Maero A., Shah R.,
 RA Saby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,
 RA Granat S., Shoddy N., Hasegawa A., Hameed A., Iodhi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana.";
 RL Nature 402:769-777(1999).
 RN [4]
 RP SEQUENCE OF 1-90 AND 118-170 FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delsen M.;
 RL Submitted (OCT-1992) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE NATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M22034; AAA32744.1; -
 DR EMBL; Z24745; CAA80871.1; -
 DR EMBL; A1035680; CAB38845.1; -
 DR EMBL; A1161566; CAB79570.1; -
 DR EMBL; Z17598; CAA75010.1; -
 DR EMBL; Z17594; CAA79008.1; -
 DR PIR; JAO162; NMMU2.
 DR PIR; S34677; S34677.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001768; Cereal_tryp_aml_1nh.
 DR InterPro; IPR000617; Napin.
 DR Pfam; PF00234; tryp_alpha_aml1; 1.
 DR PRINTS; PR00496; NAPIN.
 DR ProDom; PD002498; Napin; 1.
 DR SMART; SM00499; AAI; 1.
 KW Seed storage protein; Albumin; Signal; Multigene family.
 FT SIGNAL 1 21
 FT PROPEP 22 37
 FT CHAIN 38 72
 FT PROPEP 73 88
 FT CHAIN 89 170
 FT PROPEP 73 88
 FT CHAIN 89 170
 SQ SEQUENCE 170 AA; 19361 MW; 0A0562D5FAB56275 CRC64;
 2S52 ARATH Length: 170 September 12, 2002 12:09 Type: P Check: 3182 ..
 Found using 'consensus' (consens.key)

 1 MANKFLVCATFALCFLTNASTYRTVEFEDDASNPMGPRQCKQCFQSOHLRACOK 45

 61 LMRKMGKGGKGRSLDDERFDEDDINPGQPGQGHIIQQCCSELKQREPCVCPTLRO

 121 AARAVSLGGHGFQSGSRKRYKTAKYLPNICKIQGVGCPQTITPFFPPY 158

 1 match found in sequence:
 2S52capma; Mabinlin II precursor (MAB II) (Sweet protein).
 (from 'consensus.sp.pep')
 TOIG of: 2S52_capma check: 6217 from: 1 to: 155

 ID 2S52_CAPMA STANDARD; PRT; 155 AA.
 AC P30233; O04774;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mabinlin II precursor (MAB II) (Sweet protein).
 OS Capparis masekai (Mabinlang).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Caparis.
 OC NCBI_TaxID=13395;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Seed;
 RX MEDLINE=97128796; PubMed=8973336;
 RA Nirasawa S., Masuda Y., Nakaya K., Kurihara Y.;
 RT "Cloning and sequencing of a cDNA encoding a heat-stable sweet
 RT protein, mabinlin II.";
 RL Gene 181:225-227(1996).

[2]
 RN SEQUENCE OF 36-68 AND 83-154.
 RP TISSUE=Seed;
 RC MEDLINE=93145958; PubMed=8425538;
 RA Liu X., Maeda S., Hu Z., Aluchi T., Nakaya K., Kurihara Y.;
 RT "Purification, complete amino acid sequence and structural
 RT characterization of the heat-stable sweet protein, mabinlin II.";
 RN Eur. J. Biochem. 211:281-287(1993).
 RN [3]
 RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
 RC MEDLINE=94002261; PubMed=8393931;
 RA Nirasawa S., Liu X., Nishino T., Kurihara Y.;
 RT "Disulfide bridge structure of the heat-stable sweet protein mabinlin
 RT II.";
 RL Biochim. Biophys. Acta 1202:277-280(1993).
 CC -1- FUNCTION: HEAT STABLE 2S SEED STORAGE PROTEIN HAVING SWEETNESS-
 CC INDUCING ACTIVITY.
 CC -1- SUBUNIT: HETERODIMER OF A SMALL A AND A LARGE B CHAIN LINKED BY
 CC DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D83997; BAA12204.1; -;
 DR Mendel; 14634; Capma; 1175; 14634.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR000617; Napin.
 DR Pfam: PF01631; Seedstore_2S; 1.
 DR PRINTS; PR00496; NAPIN.
 DR PRODom; PD002498; Napin; 1.
 DR SMART; SM00499; AAI; 1.
 DR Seed storage protein: Albumin; Signal: Sweet-taste.
 KW SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 35
 FT CHAIN 36 68 MABINLIN II, A CHAIN.
 FT PROPEP 69 82
 FT CHAIN 83 154 MABINLIN II, B CHAIN.
 FT PROPEP 155 155
 FT DISULFID 40 103
 FT DISULFID 53 92
 FT DISULFID 93 141
 FT DISULFID 105 149
 FT MOD_RES 36 36 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 83 83 PYRROLIDONE CARBOXYLIC ACID.
 FT CONFLICT 148 148 A -> T (IN REF. 1).
 FT CONFLICT 153 153 A -> T (IN REF. 1).
 SQ SEQUENCE 155 AA; 18089 MW; 72E885DECC2D46A CRC64;
 2SS3_CAPMA Length: 155 September 12, 2002 12:09 Type: P Check: 6217
 Found using 'consensus' (consens.key)

 1 MAKILFATLALFVLLANMSIQTVIEVDEEDNOLMRCCROFLQHORLRACORFIHRR
 40
 61 ADFGGQDELEDEVEDNDENQRRRALRCCNQQLQVDRPCVPLRQAQVYLRQI

 121 IOGPQLRLFDARNLPNICININIGACPRAMP
 149

 1 match found in sequence:
 2ss3arath : 2s seed storage protein 3 precursor (2s albumin storage protein)
 (from "consensus_sp.pep")

TOTG of: 2ss3.arath check: 419 from: 1 to: 164
 ID 2SS3_ARATH STANDARD; PRT; 164 AA.
 AC P15459;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 2S seed storage protein 3 precursor (2S albumin storage protein)
 DE (NMW02-2S albumin 3).
 DE AT2S3 OR AT4G27160 OR T24A18.110.
 GN Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 CX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA van Damme J., Segura M., Cheysen G., van Montagu M.,
 RA Vandekerckhove J.,
 RT "Determination of the processing sites of an Arabidopsis 2S albumin
 RT and characterization of the complete gene family.";
 RL Plant Physiol. 87:859-866(1988).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. C24;
 RA Conceicao A.D.S., Krebbers E.,
 RA Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RX MEDLINE=20083488; PubMed=10617198;
 RA Mayer K.F.X., Schueller C., Wambutt R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entian K.-D., Terryn N.,
 RA Harris B., Ansoorge W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermair B., Mache R., Meller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidheini T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohnselt J., Zimmermann W., Wedler H., Ridley P.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA Van der Schueren J., Gymonprez B., Chuang Y.-J., Vandenbussche F.,
 RA Braeken M., Welfjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Wiltzenger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dikse W.,
 RA Mooijman P., Klein Lankhorst R., Rose M., Haut J., Koeltter P.,
 RA Berneriser S., Hempel S., Feldpausch M., Lambrecht S., Van den Daele H.,
 RA de Keyser A., Buysschaert C., Gieles J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., Melay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Bloeker H., Scharfe M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Grandjean R., Daurer D., Herzl A.,
 RA Neumann S., Argitrou A., Vitale D., Liguori R., Piravandi F.,
 RA Massenet O., Quigley F., Clabaud G., Wendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chetdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Barques M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Jacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Biele C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bavan M., Wilson K.R., de la Bastide M., Haeremans K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Wards E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Yordan C.,
 RA Ma P., Zhong J., Preston R., Vili D., Shekher M., Matero A., Shah R.,

RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lodhi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.,
 RT "Sequence and analysis of chromosome 4 of the plant *Arabidopsis*
 RT *thaliana*." Nature 402:769-777(1999).
 RL Nature 402:769-777(1999).
 RN [4]
 RP SEQUENCE OF 103-164 FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Raynal M., Grellet F., Laudie M., Meyer Y., Cooke R., Delzeny M.,
 RL Submitted (Oct-1992) to the EMBL/Genbank/DBD databases.
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 CC LINKED BY DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M22035; AAA22745.1; -
 DR EMBL: Z24744; CAAB0868.1; -
 DR EMBL: AL035680; CAB38846.1; -
 DR EMBL: AL161566; CAB79571.1; -
 DR EMBL: Z17580; CAA79001.1; -
 DR PIR: JAO163; NMN03.
 DR PIR: S34674; S34674.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereal_1_1.
 DR InterPro: IPR000617; Napin.
 DR Pfam: PF00234; tryp_alpha_1; 1.
 DR PRINTS: PR00496; NAPIN.
 DR ProDom: PD002498; Napin; 1.
 DR SMART: SM00499; AAI; 1.
 DR Seed storage protein; Albumin; Signal; Multigene family.
 KW SIGNAL 1 21
 FT PROPEP 22 37
 FT CHAIN 38 72 2S SEED STORAGE PROTEIN 3 SMALL SUBUNIT
 FT (BY SIMILARITY).
 FT PROPEP 73 81
 FT CHAIN 82 164 2S SEED STORAGE PROTEIN 3 LARGE SUBUNIT
 FT (BY SIMILARITY).
 SQ SEQUENCE 164 AA: 18762 MW: C9BBB6718549F248 CRC64;
 2SS3 ARATH length: 164 September 12, 2002 12:09 Type: P Check: 419
 Found using 'consensus' (consens.key)
 1 MANKFLVCAATLALCFLLTNASTIRTVVEFEEDDASNPVGPORCRCKEQSOHLRACOR
 45
 61 WMSKMRGGRGGPSLDDEPDEFGPGQGYQLAQCCNELRQEPVVCPTLKAANAVSL
 121 QGGHGPFQSKRIYQSAKYLPIKICKIQGVGCPQTITPPFPY
 151
 1 match found in sequence:
 2ss3capma; Mabinlin IIV, A and B chains (MAB IIV) (Sweet protein).
 (from "consensus-sp.pep")
 TOIG of: 2ss3-capma check: 3177 from: 1 to: 104
 ID 2SS3_CAPMA STANDARD: PRT; 104 AA.
 AC P80352;
 DT 01-FEB-1995 (Rel. 31, Created)
 RT 01-FEB-1995 (Rel. 31, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mabinlin IIV, A and B chains (MAB IIV) (Sweet protein).
 OS Capparis masaiikai (Mabinlang).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Capparidaceae;
 OC NCBI_TaxID=13393;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seed;
 RX MEDLINE=94333405; PubMed=8055976;
 RA Nirasawa S., Nishino T., Katahira M., Uesugi S., Hu Z., Kurihara Y.,
 RT "Structures of heat-stable and unstable homologues of the sweet
 RT protein mabinlin. The difference in the heat stability is due to
 RT replacement of a single amino acid residue."
 RL Eur. J. Biochem. 223:989-995(1994).
 CC -1- FUNCTION: HEAT STABLE 2S SEED STORAGE PROTEIN HAVING SWEETNESS-
 CC INDICING ACTIVITY.
 CC -1- SUBUNIT: HETERODIMER OF A SMALL A AND A LARGE B CHAIN LINKED BY
 CC DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC InterPro: IPR003612; AAI.
 CC InterPro: IPR000617; Napin.
 CC Pfam: PF01631; Seedstore_2s; 1.
 DR PRINTS: PR00496; NAPIN.
 DR ProDom: PD002498; Napin; 1.
 DR SMART: SM00499; AAI; 1.
 DR Seed storage protein; Albumin; Sweet-taste.
 KW Seed storage protein; Albumin; A CHAIN.
 FT CHAIN 1 32
 FT NON CONS 32 33
 FT CHAIN 33 104
 FT DISULFID 4 53
 FT DISULFID 17 42
 FT DISULFID 43 91
 FT DISULFID 55 99
 SQ SEQUENCE 104 AA: 12284 MW: 102EE3F5F24AD3D0 CRC64;
 2SS3_CAPMA length: 104 September 12, 2002 12:09 Type: P Check: 3177
 Found using 'consensus' (consens.key)
 1 EPLCRQFOOHGLRACQRYLRRRAQRGLADEQRGPAALRLCNOIRQVKKPCVPLRQ
 4
 61 AAHQLYGQIEGPROVRLFRANLPIKICKIPAVGRQFTRM
 99
 1 match found in sequence:
 2ss4capma; Mabinlin IIV, A and B chains (MAB IIV) (Sweet protein).
 (from "consensus-sp.pep")
 TOIG of: 2ss4-capma check: 9891 from: 1 to: 100
 ID 2SS4_CAPMA STANDARD: PRT; 100 AA.
 AC P80353;
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mabinlin IIV, A and B chains (MAB IIV) (Sweet protein).
 OS Capparis masaiikai (Mabinlang).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Capparidaceae;
 OC NCBI_TaxID=13393;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Seed;
 RX MEDLINE=94333405; PubMed=8055976;
 RA Nirasawa S., Nishino T., Katahira M., Uesugi S., Hu Z., Kurihara Y.,
 RT "Structures of heat-stable and unstable homologues of the sweet
 RT protein mabinlin. The difference in the heat stability is due to
 RT replacement of a single amino acid residue."
 RL Eur. J. Biochem. 223:989-995(1994).
 CC -1- FUNCTION: HEAT STABLE 2S SEED STORAGE PROTEIN HAVING SWEETNESS-
 CC INDICING ACTIVITY.
 CC -1- SUBUNIT: HETERODIMER OF A SMALL A AND A LARGE B CHAIN LINKED BY
 CC DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC InterPro: IPR003612; AAI.
 CC InterPro: IPR000617; Napin.
 CC Pfam: PF01631; Seedstore_2s; 1.
 DR PRINTS: PR00496; NAPIN.
 DR ProDom: PD002498; Napin; 1.
 DR SMART: SM00499; AAI; 1.
 DR Seed storage protein; Albumin; Sweet-taste.
 KW Seed storage protein; Albumin; A CHAIN.
 FT CHAIN 1 32
 FT NON CONS 32 33
 FT CHAIN 33 104
 FT DISULFID 4 53
 FT DISULFID 17 42
 FT DISULFID 43 91
 FT DISULFID 55 99
 SQ SEQUENCE 104 AA: 12284 MW: 102EE3F5F24AD3D0 CRC64;
 2SS3_CAPMA length: 104 September 12, 2002 12:09 Type: P Check: 3177
 Found using 'consensus' (consens.key)

```

RL      Eur. J. Biochem. 223:989-995(1994) .
CC      -I- FUNCTION: HEAT STABLE 2S SEED STORAGE PROTEIN HAVING SWEETINSS-
CC      INDUCING ACTIVITY
CC      -I- SUBUNIT: HETERODIMER OF A SMALL A AND A LARGE B CHAIN LINKED BY
CC      DISULFIDE BONDS.
CC      -I- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
DR      InterPro: IPR003612; AAI.
DR      InterPro: IPR000617; NapIn.
DR      Pfam: PF01631; Seedstore_2S; 1.
DR      PRINTS; PR002496; NapIN.
DR      ProDom; PD002496; NapIn; 1.
DR      SMART; SM00499; AAI; 1.
KW      Seed storage protein; Albumin; Sweet-taste.
FT      CHAIN      1      28      MABINLIN IV, A CHAIN.
FT      NON CONS   28      29
FT      CHAIN      29      100      MABINLIN IV, B CHAIN.
FT      DISULFID   4      49      BY SIMILARITY.
FT      DISULFID   17     38      BY SIMILARITY.
FT      DISULFID   39     87      BY SIMILARITY.
FT      DISULFID   51     95      BY SIMILARITY.
SQ      SEQUENCE   100 AA;  11928 MW;  7B09673EAB7793CA CRC64;
2S5A_CAFMA Length: 100 September 12, 2002 12:09 Type: P Check: 9891
Found using 'consensus' (consens.key)

```

```

1  EPICRRQFOQHILRACQRYLRRRARGDGRGALPLCCNQLRYQVKPCVCPYLROAHQ
4
-----|
61 QLYQGIEGPROVRLFRARNLPNICKIPANGRCQFTW
95

```

2 matches found in sequence:
2ss5fhehan ; 2S seed storage protein precursor (2S albumin storage protein)
(from "consensus_sp.pep")
TOIG of: 2ss5_hehan check: 2705 from: 1 to: 295

ID	2SS55_HELAN	STANDARD;	PRT;	295 AA.
AC	P15461;			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	01-OCT-1994 (Rel. 30, Last annotation update)			
DE	2S seed storage protein precursor (2S albumin storage protein).			
GN	HAG5.			
OS	Helianthus annuus (Common sunflower).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;			
OC	Asteridae: euasterids II; Asterales; Asteraceae; Asteroideae;			
OC	Heliantheae; Helianthus.			
OX	NCBI_TaxID=4232;			
XP	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 162-173.			
RC	STRAIN=CV GIANT GREY STRIPE.			
RX	MEDLINE=8142538; PubMed=2830455;			
RA	Allen R.D., Cohen E.A., Vonder Haar R.A., Adams C.A., Ma D.P.,			
RA	Nessler C.L., Thomas T.L.;			
RT	"Sequence and expression of a gene encoding an albumin storage			
RT	protein in sunflower.";			
RL	Mol. Gen. Genet. 210:211-218(1987).			
CC	-1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.			
CC	-1- PM: IT IS POSSIBLE THAT THE 38 KDa PRECURSOR IS CLEAVED INTO TWO			
CC	POLYPEPTIDES THAT ARE APPROXIMATELY THE SAME SIZE. THE MATURE			
CC	PROTEIN IS COMPOSED OF A SINGLE POLYPEPTIDE CONTAINING ONE OR MORE			
CC	INTRA-MOLECULAR DISULFIDE LINKAGES.			
CC	-1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC  entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X06410; CAA29699.1; -.
DR  PIR; S01062; S01062.
DR  InterPro; IPR003612; AAI.
DR  InterPro; IPR001768; Cereal_1rrp_amy1_inh.
DR  Pfam; PF00234; 1rrp_alpha_amy1_2.
DR  SMART; SM00499; AAI; 2.
KW  Seed storage protein; Albumin; Signal; Multigene family.
FT  SIGNAL 1 20 POTENTIAL.
FT  PROPEP 21 161 POTENTIAL.
FT  CHAIN 162 295 2S SEED STORAGE PROTEIN.
SQ  SEQUENCE 295 AA; 34071 MW; 8958BA106805142A1 CRC64;
2SS5..HELAN Length: 295 September 12, 2009 12:09 Type: P Check: 2705
Found using 'consensus' (consens.key)

```

128 QOLKQALPNVCNLQSRCEIGTITTTVESNIDI FRDRPFGTGSQOCETEIEQRP 178
178 178

188 GECQRFVEQMQSPKSTRPYQQRPGQQQQRGLQQCCNELQNVKRECHCEAIQEVAI

248 RVMRPPQOQQQQRRGQFGQEMETARRVIONLPNQCDLEVQQCTTCTG
290
293

```

1 match found in sequence:
2ssberex ; 2S sulfur-rich seed storage protein (Allergen Ber e 1)
   (from "consensus_sp.pep")
   TOIG of: 2ss_berex check: 1431 from: 1 to: 146

ID      2SS_BEREX      STANDARD;      PRT;      146 AA.
AC      P04403; P04402;
DT      20-MAR-1987 (Rel. 04, Created)
DT      01-OCT-1989 (Rel. 12, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      2S sulfur-rich seed storage protein precursor (Allergen Ber e 1).
GN      BE2S1 AND BE2S2.
OS      Bertholletia excelsa (Brazil nut).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots;
OC      Asteridae; Ericales; Lecythidaceae; Bertholletia.
OX      NCBI_TaxID=3645;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Altenbach S.B., Pearson K.W., Leung F.W., Sun S.S.M.;
RT      "Cloning and sequence analysis of a cDNA encoding a Brazil nut protein
RT      exceptionally rich in methionine.";
RL      Plant Mol. Biol. 8:239-250(1987).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Passauer R.;
RN      Submitted (DEC-1990) to the EMBL/GenBank/DBJ databases.
RL      [3]

```

RX MEDLINE=91370890; PubMed=1840683;
 RA Gander E.S., Holmstroem K.O., de Paiva G.R., de Castro L.A.B.,
 RA Carneiro M., Grossi de Sa M.F.;
 RT "Isolation, characterization and expression of a gene coding for a 2S
 albumin from *Bertholletia excelsa* (Brazil nut).";

```

RL Plant Mol. Biol. 16:437-448(1991).
RN [4]
RP SEQUENCE OF 37-64 AND 70-142.
RX MEDLINE=87004679; PubMed=3758080;
RA Ampe C., van Damme J., de Castro L.A.B., Sampaio M.J.A.M.,
  van Montagu M., Vandekerckhove J.;
RT "The amino-acid sequence of the 2S sulphur-rich proteins from seeds
  of Brazil nut (Bertholletia excelsa H.B.K.).";
RL Eur. J. Biochem. 159:597-604(1986).
CC CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
CC CC LINKED BY DISULFIDE BONDS.
CC CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isd-sib.ch/announce/
CC CC or send an email to license@sib-slb.ch).
-----
DR EMBL: M17146; AAA33010.1; -
DR EMBL: X57027; CAA40343.1; -
DR EMBL: X57028; CAA40344.1; -
DR EMBL: X54490; CAA38362.1; -
DR EMBL: X54491; CAA38363.1; ALT-SEQ.
DR EMBL: A13818; CAA01131.1; -
DR PIR: S06252; S06252.
DR PIR: A25802; A25802.
DR PIR: B25802; B25802.
DR PIR: S14946; S14946.
DR PIR: S14479; S14479.
DR PIR: S21640; S21640.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_1tryp_1.
DR Pfam: PF00234; 1tryp_alpha_1.
DR SMART: SM00499; AAI; 1.
KW Seed storage protein; Signal; Allergen.
FT SIGNAL 1 22
FT PROPEP 23 36
FT CHAIN 37 64 SMALL CHAIN.
FT PROPEP 65 69
FT CHAIN 70 142 LARGE CHAIN 1B.
FT PROPEP 143 146
FT MOD_RES 37 37 PYRROLIDONE CARBOXYLIC ACID.
FT VARIANT 91 91 S -> E (IN VARIANT 1A).
FT CONFLICT 38 39 EE -> QQ (IN REF. 4).
FT CONFLICT 122 122 I -> M (IN REF. 4).
FT CONFLICT 126 126 I -> L (IN REF. 4).
SQ SEQUENCE 146 AA; 16911 MW; A7DF77BFD76410D CRC64;

2SS_BEREX Length: 146 September 12, 2002 12:09 Type: P Check: 1431
Found using 'consensus' (consens.key)

1 MAKISVAAALLVLMALGHATAFRATVTTTVVBEENQECRQMRQOMLSHCRTMYMRQO
40

61 MESSPYQTPRRRCMEPHMSECCQQLGMDSCRCGLRMMRMRRMQEEMQPRGEGMRRM
-----
121 RLAEPTSRCLNSPMRCPMGGSIAG
137

2 matches found in sequence:
2ssricco ; 2S albumin precursor (Allergen Ric c 1).
(from "consensus.sp.pep")
TOIG of: 2ss_ricco check: 1269 from: 1 to: 258

```

ID	25S_RICCO	STANDARD:	PRT:	258 AA.
AC	P01089:			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	2S albumin precursor (Allergen Ric c 1).			
OS	Ricinus communis (Castor bean).			
OC	Euphorbia: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucots 1; Malpighiales; Euphorbiaceae; Ricinus.			
OX	NCBI_TaxID=3988;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Endosperm;			
RX	MEDLINE=91109729; PubMed=2274038;			
RA	Irvine S.D., Keen J.N., Findlay J.B.C., Lord J.M.;			
RT	"The Ricinus communis 2S albumin precursor: a single preproprotein may be processed into two different heterodimeric storage proteins.";			
RT	Mol. Gen. Genet. 222:400-408(1990).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=Endosperm;			
RX	MEDLINE=91016940; PubMed=2216785;			
RA	Irvine S.D., Lord J.M.;			
RT	"Nucleotide sequence of a Ricinus communis 2S albumin precursor gene.";			
RT	Nucleic Acids Res. 18:5890-5890(1990).			
RN	[3]			
RP	SEQUENCE OF 157-190 AND 194-258.			
RX	MEDLINE=83082772; PubMed=7174664;			
RA	Sharief F.S., Li S.S.-L.;			
RT	"Amino acid sequence of small and large subunits of seed storage protein from Ricinus communis.";			
RT	J. Biol. Chem. 257:14753-14759(1982).			
RN	[4]			
RP	SIMILARITY TO PROTEINASE INHIBITORS.			
RX	MEDLINE=83308577; PubMed=6615448;			
RA	Odani S., Koide T., Ono T., Onishi K.;			
RT	"Structural relationship between barley (Hordeum vulgare) trypsin inhibitor and castor-bean (Ricinus communis) storage protein.";			
RL	Biochem. J. 213:543-545(1983).			
CC	-1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.			
CC	-1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN LINKED BY 2 DISULFIDE BONDS.			
CC	-1- PTM: THE TWO DISULFIDE BONDS LINKING THE SMALL AND THE LARGE CHAINS INVOLVE CYS-162 AND CYS-175.			
CC	-1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)			
CC	-----			
DR	EMBL; X54158; CAA38097.1; .			
DR	PIR; A01328; RZCS.			
DR	PIR; S11499; S11499.			
DR	PIR; S11500; S11500.			
DR	PIR; S11501; S11501.			
DR	InterPro; IPR003612; AAI.			
DR	InterPro; IPR001768; Cereal_tryp_amy1Inh.			
DR	Pfam; PF00234; tryp_alpha_amy1_2.			
DR	SMART; SM00499; AAI; 2.			
DR	Seed storage protein: Signal; Allergen.			
KW	SIGNAL	1	21	PROBABLE.
FT	PROPEP	22	156	
FT	CHAIN	157	190	2S ALBUMIN, SMALL CHAIN.
FT	PROPEP	191	193	
FT	CHAIN	194	258	2S ALBUMIN, LARGE CHAIN.
FT	MOD_RES	194	194	PYRROLIDINE CARBOXYLIC ACID.
FT	CONFLICT	222	222	E -> Q (IN REF. 3).

FT CONFLICT 226 229 MISSING (IN REF. 3).
 FT CONFLICT 234 234 D -> N (IN REF. 3).
 FT CONFLICT 255 255 E -> Q (IN REF. 3).
 SO SEQUENCE 258 AA: 29290 MW; 27874CFG50E41072 CRC64;
 258_RICCO Length: 258 September 12, 2002 12:09 Type: P Check: 1269
 Found using 'consensus' (consens.key)

1 MAKLIPTALVSVLLFTIANASFAVTRTTTTEIDESKGERBSSSSQCCROEVQRKDLSS
 49

61 CERYLRSSSRSPGEVLRMPGDENQOQESQLOQCCNOYKQVDEQCEAKIYIADQ

121 IOOGLHGESEFVAVQAGEIVSSCGVRCMROTFTNPQCGCQIQBQNLQCCQETIK
 149 162

181 QQVSGQPRRSNDQERSLGGCDHLKQMSQCRGRLQAIHQOOSQOQLOGQDVFEAFR

241 TAANLPSMCGVSPTECRF
 256

1 match found in sequence:
 258bra; Napin IA and IB small chain and large chains.
 (from 'consensus.sp.pep')
 TOIG of: 2581_bra; check: 5085 from: 1 to: 110

ID 2581_BRANA STANDARD; PRT; 110 AA.
 AC P24565;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-MAR-1992 (Rel. 21, Last annotation update)
 DE Napin IA and IB small chain and large chains.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Brassica.
 NCBI_TaxID=3708;
 [1]

RP SEQUENCE.
 RC TISSUE=Seed;
 RX MEDLINE=92111741; PubMed=1765156;
 RA Monsalve R.I., Lopez-Otin C., Villalba M., Rodriguez R.;
 RT "A new distinct group of 2 S albumins from rapeseed. Amino acid
 sequence of two low molecular weight napins.";
 RL FEBS Lett. 295:207-210(1991).
 CC -1- FUNCTION: THE SMALL, BASIC, WATER-SOLUBLE NAPINS ARE ONE OF THE
 TWO MAJOR KINDS OF STORAGE PROTEINS SYNTHESIZED IN THE SEED DURING
 ITS MATURATION.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 LINKED BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: NAPIN IA AND IB ARE MINOR COMPONENT OF SEED 2S
 ALBUMIN.
 CC -1- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF NAPIN IA.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC PIR: S20350; S20350.
 DR InterPro: IPR000617; AAI.
 DR InterPro: IPR000617; Napin.
 DR Pfam: PF01631; Seedstore_2S; 1.
 DR PRINTS: PR00496; NAPIN.
 DR PRODOM: PD002498; Napin; 1.
 DR SMART: SM00499; AAI; 1.
 KW Seed storage protein.
 FT CHAIN 1 31 SMALL CHAIN.
 FT NON_CONS 31 32
 FT CHAIN 32 110 LARGE CHAIN.

FT VARIANT 31 31 MISSING (IN MINOR FORM "SM", LESS
 FT VARIANT 37 37 THAN 7%).
 FT VARIANT 107 110 E -> Q (IN NAPIN IB).
 SO SEQUENCE 110 AA: 12691 MW; 3A2938ADA2C1E995 CRC64;
 2581_BRANA Length: 110 September 12, 2002 12:09 Type: P Check: 5085
 Found using 'consensus' (consens.key)

1 QPKQREFOEQHILRACQWILRQAGSPQSPQEGPWLREDCNBLYOEDQVCVPT
 5

61 LKQAKSVYVQGGHGFQSTRIYQIAKNLPNVCNMKQIGTCFPIAPFP
 101

1 match found in sequence:
 allbraju; Allergen Bra j 1-E, small and large chains (Bra j I).
 (from 'consensus.sp.pep')
 TOIG of: all_braju check: 4076 from: 1 to: 129

ID ALL1_BRAJU STANDARD; PRT; 129 AA.
 AC P80215;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Allergen Bra j 1-E, small and large chains (Bra j I).
 OS Brassica juncea (Leaf mustard) (Indian mustard).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Brassica.
 NCBI_TaxID=3707;
 [1]
 RN SEQUENCE.
 RP TISSUE=Seed;
 RC MEDLINE=93356721; PubMed=7689955;
 RX Monsalve R.I., Gonzalez de la Pena M.A., Menendez-Arias L.,
 RA Lopez-Otin C., Villalba M., Rodriguez R.;
 RT "Characterization of a new oriental-mustard (Brassica juncea)
 Allergen, Bra j IE: detection of an allergenic epitope.";
 RL Biochem. J. 293:625-632(1993).
 CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
 CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
 LINKED BY TWO DISULFIDE BONDS.
 CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
 CC InterPro: IPR000617; AAI.
 DR InterPro: IPR000617; Napin.
 DR Pfam: PF01631; Seedstore_2S; 1.
 DR PRINTS: PR00496; NAPIN.
 DR PRODOM: PD002498; Napin; 1.
 DR SMART: SM00499; AAI; 1.
 KW Allergen; Seed storage protein.
 FT CHAIN 1 37 SMALL CHAIN.
 FT NON_CONS 37 38
 FT CHAIN 38 129 LARGE CHAIN.
 FT VARIANT 6 6 F -> I.
 FT VARIANT 20 20 R -> K.
 SO SEQUENCE 129 AA: 14644 MW; D6F28B03P62B08P8 CRC64;
 ALL1_BRAJU Length: 129 September 12, 2002 12:09 Type: P Check: 4076
 Found using 'consensus' (consens.key)

1 AGPRFPKRCRKEFOQAOHLRACQWILRQAGSGPQSPQGPPLLOCCNELHQEE
 9

61 PLVCPTLKASKAVKQOIRQGGQGGQGLQHEISRITQTATHLPRVNCIRVSTCP
 119
 121 FQKTPGSPS

```

-----
1 match found in sequence:
all:signal : Allergen Sin a 1, small and large chains (Sin a I).
(from "consensus.sp.pep")
TOIG of: all:signal check: 2875 from: 1 to: 127

ID ALL-SIGNAL STANDARD; PRT; 127 AA.
AC P15322;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
OS Allergen Sin a 1, small and large chains (Sin a I).
OS Sinapis alba (White mustard) (Brassica hirta).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eustoids II; Brassicales; Brassicaceae; Sinapis.
OX NCBI_TaxID=3728;
RN [1]
RP SEQUENCE.
RC TISSUE=Seed;
RX MEDLINE=89030681; PubMed=3181153;
RA Mendez-Arias L., Moneo I., Dominguez J., Rodriguez R.;
RT "Primary structure of the major allergen of yellow mustard (Sinapis
RL alba L.) seed, Sin a I.";
RL Eur. J. Biochem. 177:159-166(1988).
CC -1- FUNCTION: THIS IS A 2S SEED STORAGE PROTEIN.
CC -1- SUBUNIT: THE PROTEIN CONSISTS OF TWO CHAINS LINKED BY DISULFIDE
CC BONDS.
CC -1- DISEASE: ALLERGIC DISEASE, CABBAGE ALLERGY.
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
DR PIR; S01792; S01792.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000617; Napin.
DR Pfam; PF01631; Seedstore_2S; 1.
DR PRINTS; PR00496; NAPIN.
DR PRODOM; PD002498; Napin; 1.
DR SMART; SM00499; AAI; 1.
KW Allergen; Seed storage protein.
FT CHAIN 1 39 SMALL CHAIN.
FT NON_CONS 39 40
FT CHAIN 40 127 LARGE CHAIN.
FT VARIAM 6 6 R->G.
SQ SEQUENCE 127 AA; 14180 MW; 4CD920284F04EEFD CRC64;

ALL-SIGNAL Length: 127 September 12, 2002 12:09 Type: P Check: 2875 ..
Found using 'consensus' (consens.key)

1 PACGFRIPKCKRFFQAOHLRAQQMLHKQAMQSGSGSPGQPPQLQCCNELHQE
10
-----
61 EPLVCPTLKASRAVQOVROQLQOQCGPHVISRIYQATHLPKVCNIPQVSCPFK
117
-----
121 KTMGPPS
-----
1 match found in sequence:
g119orysa : 19 kda globulin precursor (Alpha-globulin).
(from "consensus.sp.pep")
TOIG of: g119_orysa check: 1252 from: 1 to: 186

ID GL19-ORYSA STANDARD; PRT; 186 AA.
AC P29835;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE 19 kda globulin precursor (Alpha-globulin).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

```

```

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=CV, NIPPONBARE; TISSUE=Endosperm.
RX MEDLINE=92119226; PubMed=1731968;
RA Shirosh B.S., Wen L., Zen K.C., Huang J.K., Pan J.S., Hermodson M.A.,
RA Tanaka K., Muthukrishnan S., Reeck G.R.;
RT "A novel cereal storage protein: molecular genetics of the 19 kda
RL globulin of rice.";
RN Plant Mol. Biol. 18:151-154(1992).
RN [2]
RP SEQUENCE OF 5-186 FROM N.A.
RC STRAIN=CV, LAMONT; TISSUE=Endosperm;
RX MEDLINE=93277591; PubMed=8503935;
RA Krishnan H.B., Pueppke S.G.;
RT "Nucleotide sequence of an abundant rice seed globulin: homology with
RL the high molecular weight glutelins of wheat, rye and triticale.";
RL Biochem. Biophys. Res. Commun. 193:460-466(1993).
CC -1- FUNCTION: SEED STORAGE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X63990; CAA45400.1; -
DR EMBL; L12252; AAA72362.1; ALT_INTT.
DR PIR; S20024; WMR219.
DR HSSP; P01085; IHSS.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_aml_1nh.
DR Pfam; PF00234; tryp_alpha_aml1; 1.
DR SMART; SM00499; AAI; 1.
KW Signal; Seed storage protein.
FT SIGNAL 1 22
FT CHAIN 23 186 19 KDA GLOBULIN.
SQ SEQUENCE 186 AA; 21050 MW; 9E09BA74CBB06810 CRC64;

GL19-ORYSA Length: 186 September 12, 2002 12:09 Type: P Check: 1252 ..
Found using 'consensus' (consens.key)

1 MASKVVFPAALMAVVAISGAHVSESEMRFRDQCQREVODSPLDAGROYLDRLTGRE
36
-----
61 RQPMFRPPALGLRMQCCQQLDVSRECRCAIRRAVRYSEEMPRPLDQNGSSSSEY
-----
121 YGEGSSSEQGYEGESSEGYEGEQOQPGMTRVRLTRARQVAAQJPSMCRVPEQCCSI
178
-----
181 FAGQY
-----
1 match found in sequence:
1a03wheat : Alpha-amylase/trypsin inhibitor CM3 precursor (Chloroform/methanol
(from "consensus.sp.pep")
TOIG of: 1a03_wheat check: 7810 from: 1 to: 168

ID 1A03-WHEAT STANDARD; PRT; 168 AA.
AC P17314;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last annotation update)
DE Alpha-amylase/trypsin inhibitor CM3 precursor (Chloroform/methanol-

```

DE soluble protein CM3).
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 OX NCBI_TaxID=4565;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, CHINESE SPRING; TISSUE=Endosperm;
 RA MEDLINE=91346675; PubMed=2102861;
 RT Garcia-Maroto F., Marana C., Mena M., Garcia-Olmedo F., Carbonero P.;
 RT "Cloning of cDNA and chromosomal location of genes encoding the three
 RT types of subunits of the wheat tetrameric inhibitor of insect alpha-
 RT amylase."; Plant Mol. Biol. 14:845-853(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, AGATHE; TISSUE=Seed;
 RA Gautier M.F., Alary R., Lullien V., Joudrier P.;
 RT Submitted (JUN-1991) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 26-60.
 RC STRAIN-CV, TURGIDUM; TISSUE=Endosperm;
 RA Shewry P.R., Lafandra D., Salcedo G., Aragoncillo C.,
 RA Garcia-Olmedo F., Lew E.J.-L., Dietler M.D., Kasarda D.D.;
 RT "N-terminal amino acid sequence of chloroform/methanol-soluble
 RT proteins and albumins from endosperm of wheat, barley and related
 RT species."; FEBS Lett. 175:359-363(1984).
 RL FBS Lett. 175:359-363(1984).
 CC -1- FUNCTION: ALPHA-AMYLASE/TRYPsin INHIBITOR. IT COULD BE INVOLVED
 CC IN INSECT DEFENSE MECHANISMS.
 CC -1- SUBUNIT: SUBUNIT OF THE TETRAMERIC INHIBITOR.
 CC -1- TISSUE SPECIFICITY: DEVELOPING ENDOSPERM.
 CC -1- PTM: FIVE DISULFIDE BONDS ARE PRESENT (PROBABLE), WHICH ARE
 CC ESSENTIAL FOR THE INHIBITOR ACTIVITY.
 CC -1- MISCELLANEOUS: CM PROTEINS WOULD BE INVOLVED IN THE COOKING
 CC QUALITY OF PASTA.
 CC -1- SIMILARITY: BELONGS TO THE CEREAL TRYPsin/ALPHA-AMYLASE INHIBITOR
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X17574; CAA35597.1; -;
 DR EMBL: X61032; CAA43367.1; -;
 DR PIR: S10029; S10029.
 DR PIR: S16031; S16031.
 DR HSP: P01088; IBEA.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; CereAl_tryp_aml1_inh.
 DR Pfam: PF00234; tryp_alpha_aml1; 1.
 DR PRINTS: PR00808; AMLASEINHBR.
 DR SMART: SM00499; AAI; 1.
 DR PROSITE: PS00426; CEREAL_TRYP_AMYL_INH; 1.
 KW Serine protease inhibitor; Alpha-amylase inhibitor; Multigene family;
 KW Seed; Signal.
 FT SIGNAL 1 25
 FT CHAIN 26 168 ALPHA-AMYLASE/TRYPsin INHIBITOR CM3.
 SO SEQUENCE 168 AA; 18221 MW; A950FF272E90D23C CRC64;
 IA03_WHEAT Length: 168 September 12, 2002 12:09 Type: P Check: 7810 ..
 Found using 'consensus' (consens.key)

1 MACKSCSLLLAVALSVLAASASGCVPGVAFRTNLPHCRDYLVOOTCGFTPGSK
 29 -----

61 LPEWMTASIVSPGKPYLAKLYCCQELAEISQCCREALRYFIALPVSPQVPPDRSGNVG
 121 ESGILIDPGCPREMQMDFRLVAPGQCNLATINHVRCPPAVEQPLWT
 159 -----
 1 match found in sequence:
 Iaadhorvu; Alpha-amylase/trypsin inhibitor Cwd precursor (Chloroform/methanol
 (from "consensus-sp.pep")
 ToIG of: Iaad_horvu check: 8003 from: 1 to: 171
 ID IAAD_HORVU STANDARD; PRT; 171 AA.
 AC P11643;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Alpha-amylase/trypsin inhibitor Cwd precursor (Chloroform/methanol-
 DE soluble protein Cwd).
 GN IAT3 OR CMD2.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, ABYSSINIAN; TISSUE=Endosperm;
 RX MEDLINE=9403333; PubMed=8219088;
 RA Medina-Alcazar J., Hueros G., Carbonero P.;
 RT "Cloning of cDNA, expression, and chromosomal location of genes
 RT encoding the three types of subunits of the barley tetrameric
 RT inhibitor of insect alpha-amylase."; Plant Mol. Biol. 23:535-542(1993).
 RL Plant Mol. Biol. 23:535-542(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, NK 1558;
 RX MEDLINE=97351567; PubMed=9207849;
 RA Groset J., Alary R., Gautier M.F., Menossi M.,
 RA Martine-Izquierdo J.A., Joudrier P.;
 RT "Characterization of a barley gene coding for an alpha-amylase
 RT inhibitor subunit (Cwd protein) and analysis of its promoter in
 RT transgenic tobacco plants and in maize kernels by microprojectile
 RT bombardment."; Plant Mol. Biol. 34:331-338(1997).
 RL Plant Mol. Biol. 34:331-338(1997).
 RN [3]
 RP SEQUENCE OF 11-171 FROM N.A.
 RC STRAIN-CV, HIPROLY; TISSUE=Endosperm;
 RX MEDLINE=89000796; PubMed=3167062;
 RA Halford N.G., Morris N.A., Utwin P., Williamson M.S., Kasarda D.D.,
 RA Lew E.J.-L., Kreis M., Shewry P.R.;
 RT "Molecular cloning of the barley seed protein Cwd: a variant member
 RT of the alpha-amylase/trypsin inhibitor family of cereals."; Biochim. Biophys. Acta 950:435-440(1988).
 RL Biochim. Biophys. Acta 950:435-440(1988).
 RN [4]
 RP SEQUENCE OF 25-30.
 RC STRAIN-CV, H354-295-2-5; TISSUE=starchy endosperm;
 RX MEDLINE=94170739; PubMed=8125056;
 RA Flengsard R.;
 RT "Separation of acidic barley endosperm proteins by two-dimensional
 RT electrophoresis."; J. Biol. Chem. 268:10660-10666(1993).
 RL J. Biol. Chem. 268:10660-10666(1993).
 CC -1- FUNCTION: PART OF A COMPLEX WITH INHIBITORY ACTIVITY, BUT CMD
 CC IS INACTIVE AS A SEPARATE SUBUNIT.
 CC -1- SUBUNIT: HETEROTETRAMER OF ONE CMA, ONE CMB AND TWO CMD CHAINS.
 CC -1- TISSUE SPECIFICITY: ENDOSPERM.
 CC -1- PTM: FIVE DISULFIDE BONDS ARE PRESENT (PROBABLE), WHICH ARE
 CC ESSENTIAL FOR THE INHIBITOR ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE CEREAL TRYPsin/ALPHA-AMYLASE INHIBITOR
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X69939; CAA49557.1; -
 DR EMBL: U47641; AAB63441.1; -
 DR EMBL: X13198; CAA31585.1; -
 DR PIR: J40071; J40071.
 DR HSSP: P01088; 1BEA.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereals_1tryp-amy1_inh.
 DR Pfam: PF00234; tryp_alpha_amy1; 1.
 DR PRINTS: PR00808; AMLASEINHBR.
 DR SMART: SM00499; AAI; 1.
 DR PROSITE: PS00426; CEREAL_TRYP_AMYL_INH; 1.
 DR Serine protease inhibitor; Alpha-amyase inhibitor; Multigene family;
 KW Seed: Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 171 ALPHA-AMYLASE/TRYPSTIN INHIBITOR CMD.
 SO SEQUENCE 171 AA; 18525 MW; 8D8934FED808FED4 CRC64;
 IAAH_HORVU length: 171 September 12, 2002 12:09 Type: P Check: 8003
 Found using 'consensus' (consens.key)

1 MACKSSRLLLATVWVSVPAAAAAATDCSPGVAFPNLLGHCRDYLQOTCAVFP
 32

61 GSKLPEMTSAELNYPQPLAKLYCCQELAEIPQRCCEALRYFMALPPSPQVDPSTG

121 NVQSGIMDLPGCPREMQRDEVLVLAFGQCNLATIHNVRCFAVEQPLMI
 162

1 match found in sequence:
 laehorvu ; Trypsin inhibitor CME precursor (Chloroform/methanol-soluble prote
 (from "consensus_sp.pep")
 TOIG of: laae_horvu check: 7462 from: 1 to: 144

ID IAAE_HORVU STANDARD; PRT; 144 AA.
 AC P01086; O40038;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Trypsin inhibitor CME precursor (Chloroform/methanol-soluble protein
 DE CME).
 DE ITRI.
 GN ITRI.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90158508; PubMed=2516240;
 RA Salcedo G., Molina-Cano J.L., Garcia-Olmedo F., Carbonero P.,
 RT "The gene for trypsin inhibitor CME is regulated in trans by the lys
 RT 3a locus in the endosperm of barley (Hordeum vulgare L.).",
 RL Mol. Gen. Genet. 219:474-479(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. VILVA; TISSUE=Endosperm;
 RX MEDLINE=96027931; PubMed=7476859;
 RA Diaz I., Royo J., O'Connor A., Carbonero P.,
 RT "The promoter of the gene trf1 from barley confers a different tissue
 RT specificity in transgenic tobacco.",

RL Mol. Gen. Genet. 248:592-598(1995).
 RN [3]
 RP SEQUENCE OF 25-144.
 RX MEDLINE=8328397; PubMed=6345537;
 RA Odani S., Koide T., Ono T.,
 RT "The complete amino acid sequence of barley trypsin inhibitor.",
 RL J. Biol. Chem. 258:7998-8003(1983).
 RN [4]
 RP SEQUENCE OF 25-42.
 RC STRAIN=CV. BOWT; TISSUE=Starchy endosperm;
 RX MEDLINE=21088911; PubMed=11271408;
 RA Kristoffersen H.E., Flengsrud R.,
 RT "Separation and characterization of basic barley seed proteins.",
 RL Electrophoresis 21:3693-3700(2000).
 CC -1- PTM: FIVE DISULFIDE BONDS ARE PRESENT (PROBABLE), WHICH ARE
 CC ESSENTIAL FOR THE INHIBITOR ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE CEREAL TRYPSTIN/ALPHA-AMYLASE INHIBITOR
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X65875; CAA46705.1; -
 DR EMBL: X17302; CAA35188.1; -
 DR PIR: A01325; TTBH.
 DR PIR: J00342; J00342.
 DR HSSP: P01087; ITMO.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereals_1tryp-amy1_inh.
 DR Pfam: PF00234; tryp_alpha_amy1; 1.
 DR PRINTS: PR00808; AMLASEINHBR.
 DR SMART: SM00499; AAI; 1.
 DR PROSITE: PS00426; CEREAL_TRYP_AMYL_INH; 1.
 DR Serine protease inhibitor; Seed; Multigene family; Signal.
 KW Seed: Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 144 TRYPSTIN INHIBITOR CME.
 FT ACT_SITE 57 57 INTERACTION WITH TRYPSTIN (PROBABLE).
 FT CONFLICT 143 143 G -> A (IN REF. 1).
 FT CONFLICT 144 144 Y -> YGVVL (IN REF. 2).
 SO SEQUENCE 144 AA; 15767 MW; 55D8A59499108082 CRC64;
 IAAE_HORVU length: 144 September 12, 2002 12:09 Type: P Check: 7462
 Found using 'consensus' (consens.key)

1 MAFKYOLLASAVALMLATVATATSPGSCAPGALPBNPLRACRTVVVSQICHQPRLLT
 29

61 SDRKRCDELSAIPAYCEALRTIMGVVWQAFEGAYFDSNCPREKOTSYAANL

121 VTPQECNLGTIHSAVCPELOPGY
 137

1 match found in sequence:
 Itfmaize ; Trypsin/factor XIII inhibitor precursor (Hageman factor inhibitor)
 (from "consensus_sp.pep")
 TOIG of: Itf_maize check: 5547 from: 1 to: 155

ID ITRE_MAIZE STANDARD; PRT; 155 AA.
 AC P01088;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Trypsin/factor XIII inhibitor precursor (Hageman factor inhibitor)

```

DE (CHFI).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OC NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC SPRAIN-CV A188; TISSUE=Endosperm;
RX MEDLINE=92216060; PubMed=1538956;
RA Wen L., Huang J.K., Zen K.C., Johnson B.H., Muthukrishnan S.,
RA Mackay V., Manney T.R., Manney M., Reeck G.R.;
RT "Nucleotide sequence of a cDNA clone that encodes the maize inhibitor
RT of trypsin and activated Hageman factor.";
RL Plant Mol. Biol. 18:813-814(1992).
RN [2]
RP SEQUENCE OF 29-138.
RC TISSUE=Seed;
RX MEDLINE=84239823; PubMed=6610678;
RA Mahoney W.C., Hermodson M.A., Jones B., Powers D.D., Corfman R.S.,
RA Reeck G.R.;
RT "Amino acid sequence and secondary structural analysis of the corn
RT inhibitor of trypsin and activated Hageman factor.";
RL J. Biol. Chem. 259:8412-8416(1984).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS) OF 33-148.
RX MEDLINE=99017881; PubMed=9799488;
RA Behnke C.A., Yee V.C., Le Trong I., Pedersen L.C., Stenkamp R.E.,
RA Kim S.-S., Reeck G.R., Teller D.C.;
RT "Structural determinants of the bifunctional corn Hageman factor
RT inhibitor: X-ray crystal structure at 1.95-A resolution.";
RL Biochemistry 37:15277-15288(1998).
CC -1- FUNCTION: POTENT INHIBITOR OF MAMMALIAN TRYPSIN AND A SPECIFIC
CC INHIBITOR OF FACTOR XIIA (ACTIVATED HAGEMAN FACTOR).
CC -1- SUBUNIT: MONOMER.
CC -1- MISCELLANEOUS: THE SEQUENCE HETEROGENEITY SUGGESTS THAT 2 GENES
CC EXIST FOR THIS INHIBITOR. THE GENES MAY BE ALLELES, OR MULTIPLE
CC LOCI COULD EXIST.
CC -1- SIMILARITY: BELONGS TO THE CEREAL TRYPSIN/ALPHA-AMYLASE INHIBITOR
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X54064; CAA37998.1; -.
DR PIR: S20850; TIZML.
DR PDB: 1BEA; 12-AUG-98.
DR PDB: 1BFA; 12-AUG-98.
DR MaizEDB: 25928; -.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; CereAl_Tryp_amy1_inh.
DR Pfam: PF00234; tryp_alpha_amy1; 1.
DR PRINTS: PR00808; AMTASINBTR.
DR SMART: SM00499; AAI; 1.
DR PROSITE: PS00426; CEREAL_TRYP_AMYL_INH; 1.
KW Serine protease inhibitor; Seed; Signal; 3D-structure.
FT CHAIN 1 28
FT SIGNAL 29 138
FT PROPEP 139 155
FT ACT_SITE 62 83
FT DISULFID 34 62
FT DISULFID 48 72
FT DISULFID 57 114
FT DISULFID 73 132
FT DISULFID 85 143
FT VARIANT 117 117
FT VARIANT 49 49
FT CONFLICT 56 56
FT CONFLICT 56 56
E -> A.
R -> C (IN REF. 2).
T -> R (IN REF. 2).

```

```

FT CONFLICT 62 62 R -> RPR (IN REF. 2).
FT CONFLICT 107 107 R -> A (IN REF. 2).
FT CONFLICT 116 116 R -> A (IN REF. 2).
FT CONFLICT 120 120 R -> Q (IN REF. 2).
FT CONFLICT 135 135 A -> E (IN REF. 2).
SQ SEQUENCE 155 AA; 16302 MW; A31A8AA3BA19DE7A CRC64;
ITRF_Maize Length: 155 September 12, 2002 12:09 Type: P Check: 5547
Found using 'consensus' (consens.key)
-----
1 MASSSSSHRRLLIAAVLLSVLAASASAGTSCVPGWAIPHNPPLSCRWYVTSRTGIG
34
61 PRLPPELKRRCRELADIPAYCRCTALSIIMDCAIPPGPOALEGRLEDLPQCPREVOR
-----
121 GFAATVTEACONLATISGVAECPIILGGTMPK
143
-----
1 match found in sequence:
itrysinar; Trypsin inhibitor (TISA).
(from "consensus-sp.pep")
Total of: itrysinar check: 4709 from: 1 to: 130
ID ITRY_SINAR STANDARD; PRT; 130 AA.
AC P38057;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Trypsin inhibitor (TISA).
DE Sinapis arvensis (Charlock).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Euphorbiales; Brassicales; Brassicaceae; Sinapis.
OC NCBI_TaxID=29728;
RN [1]
RP TISSUE=Seed;
RX MEDLINE=94350545; PubMed=8070965;
RA Svendsen I.B., Nicolova D., Goshay I., Genov N.;
RT "Primary structure, spectroscopic and inhibitory properties of a two-
RT chain trypsin inhibitor from the seeds of Charlock (Sinapis arvensis
RT L), a member of the napin protein family.";
RL Int. J. Pept. Protein Res. 43:425-430(1994).
CC -1- FUNCTION: INHIBITS TRYPSIN WITH A KI OF 7 X 10(-6) M.
CC -1- SUBUNIT: THE PROTEIN CONSISTS OF TWO CHAINS LINKED BY DISULFIDE
CC BONDS.
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
CC -----
DR InterPro: IPR003612; AAI.
DR InterPro: IPR000617; Napin.
DR Pfam: PF01631; Seedstore_2s; 1.
DR PRINTS: PR00496; NAPIN.
DR PRODOM: PD002498; Napin; 1.
DR SMART: SM00499; AAI; 1.
KW Seed storage protein; Multigene family; Polymorphism.
FT CHAIN 1 39
FT NON_CONS 39 40
FT CHAIN 40 130
FT VARIANT 32 32 LARGE CHAIN.
FT VARIANT 33 32 R -> M.
FT VARIANT 53 57 MISSING (IN FORM II).
FT VARIANT 73 73 A -> S.
FT VARIANT 77 77 K -> R.
FT VARIANT 81 81 Q -> R.
FT VARIANT 87 87 H -> Q.
FT VARIANT 89 89 Q -> H.
FT VARIANT 91 91 G -> Q.
FT VARIANT 97 97 E -> M.
FT VARIANT 98 98 I -> V.
FT VARIANT 99 99 R -> S.

```

FT VARIANT 106 106 T -> K.
 FT VARIANT 123 123 N -> O.
 FT VARIANT 124 124 K -> G.
 FT VARIANT 126 126 M -> V.
 SQ SEQUENCE 130 AA; 14682 MW; EC02E4B26D180DF2 CRC64;

ITRY_SINAR Length: 130 September 12, 2002 12:09 Type: P Check: 4709
 Found using 'consensus' (consens.key)

1 PAGEFRLPRCKRKEFQAQHLRACQQLHKOARSGSGSPSGPOGPPPLQQCCNELHGE
 10

61 EPLVCVPTLKGAAKAVKQIQGGQGGQGGQGLQHEIRIRYQATHLPRVCNIPQVVC
 120

121 PFNKTMPGPS

1 match found in sequence:

putwheat : Puroindoline-B precursor.
 (from "consensus-sp.pep")

FOIG of: putwheat Check: 8339 from: 1 to: 148

ID PUIB_WHEAT STANDARD; PRT; 148 AA.

AC Q10464;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Puroindoline-B precursor.

OS Triticum aestivum (wheat).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Triticeae; Triticum.

OX NCBI_Taxid=4565;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CV. CAPITOLE; TISSUE=seed;

RX MEDLINE=94272013; Pubmed=7516201;

RA Gautier M.-F., Aleman M.-F., Guirao A., Marion D., Joudrier P.;

"Triticum aestivum puroindolines, two basic cysteine-rich seed

proteins: cDNA sequence analysis and developmental gene expression.";

Plant Mol. Biol. 25:43-57(1994).

RL [2]

RP SEQUENCE OF 30-148.

RA Biochet J.E., Kaboulou A., Compoin J.P., Marion D.;

RL (In) Bushuk W., Tkachuk R. (eds.);

RL Gluten proteins, pp.314-325, American Association of Cereal Chemists,

RL St. Paul MI (1991).

CC - FUNCTION: ACTS AS A MEMBRANOTOXIN, PROBABLY THROUGH ITS

ANTIBACTERIAL AND ANTIFUNGAL ACTIVITIES, CONTRIBUTING TO THE

DEFENSE MECHANISM OF THE PLANT AGAINST PREDATORS.

- PTM: FIVE DISULFIDE BONDS ARE PRESENT.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC EMBL; X69912; CAA49537.1; -.

DR HSSP: P07597; ILIP.

DR InterPro: IPR003612; AAI.

DR InterPro: IPR001766; Cereals_tryp_aml_1nh.

DR Pfam: PF00234; tryp_alpha_aml_1.

DR SMART: SM00499; AAI: 1.

KW Membrane; Toxin; Antibiotic; Signal.

FT SIGNAL 1 19 POTENTIAL.

FT PROPEP 20 29

FT CHAIN 30 148 PUROINDOLINE-B.

FT DOMAIN 68 73 TRP-RICH.
 SQ SEQUENCE 148 AA; 16792 MW; 327904B4EBEC2C16 CRC64;
 PUIB_WHEAT Length: 148 September 12, 2002 12:09 Type: P Check: 8339
 Found using 'consensus' (consens.key)

1 MKTLFLALLALVASTTFAGYSEVGWYNEVGGGSGSQCPQERPKLSCKDYMERCF
 40

61 MKDFVTWPTKWMKGGEHEVERCKQLSQIAPQCRCDISIRVIQRLGFLIMRGEV

121 FKQLQRAQSLPSKCMGADCKFPGYYW
 140

-- Search Statistics --

Times: CPU Total Elapsed

00:00:00.00

00:00:00.00

Number of sequences searched:

19

Number of sequence hits:

19

Number of separate matches:

21

Number of sequence hits saved:

0

> 0 <
0110 Intelligenetics
> 0 <

Quest - Quick User-directed Expression Search Tool
Release 5.4

-- Outline of search "consensus-spl" --

Selected search type is key against sequence data banks or files.
Selected scope is Sequence.

Selected sequence key from "consens.key":

consensus (AA) ID consensus AA preliminary pattern
followed by
1 c
2 any character from 9 to 13 times
3 c
4 q or e or r or k
5 any character
6 a or f or g or l or m or p or t or v or w or y
7 l or l or v or m
8 any character from 15 to 39 times
9 cc
10 c or d or e or h or k or n or q or r or s
11 q or e or h
12 l or v
13 any character
14 any character
15 any character
16 any character
17 any character
18 c
19 any character
20 c
21 any character from 15 to 26 times
22 g
23 any character from 13 to 56 times
24 l or l
25 any character from 15 to 26 times
26 c
27 any character
28 v or l or l or m
29 any character from 1 to 8 times
30 c

Selected files:

File : consensus_spl.pep

-- Output Parameters --

Format Options: File Options:
Nucleic acid code matching Exact Indirect file NO
Find non-matching hits only No Sequence or key file NO
Report key used Yes List of hits Yes
Note position of hit Yes Hit display Yes
Display full annotations Yes Name and annotations Yes
Sequence context 50

-- Run Parameters --

Run mode Batch
Time to start comparison now
Notify at end of run NO

1 match found in sequence:
023982 ; CMD SUBUNIT OF TETRAMERIC ALPHA-AMYLASE INHIBITOR PRECURSOR.
(from "consensus_spl.pep")
TOIG of: 023982 check: 8141 from: 1 to: 171

ID 023982 PRELIMINARY; PRT; 171 AA.
AC 023982;
DT 01-JAN-1998 (TREMUREL.05, Created)
DT 01-JAN-1998 (TREMUREL.05, Last sequence update)
DT 01-DEC-2001 (TREMUREL.19, Last annotation update)
DE CMD SUBUNIT OF TETRAMERIC ALPHA-AMYLASE INHIBITOR PRECURSOR.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NK 1558; TISSUE=SEEDLING;
RX MEDLINE=97351567; PubMed=9207849;
RA Grosseil J., Alary R., Gautier M.F., Menossi M.,
RT Martinez-Izquierdo J.A., Joudrier P.;
RT "Characterization of a barley gene coding for an alpha-amylase
inhibitor subunit (CMD protein) and analysis of its promoter in
transgenic tobacco plants and in maize kernels by microprojectile
bombardment.";
RL Plant Mol. Biol. 34:331-338(1997).
DR EMBL; X69911; CAA49536.1; -
DR HSSP; P01088; IBEA.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_1tryp_1inh.
DR Pfam; PF00234; tryp_alpha_1.
DR PRINTS; PR00808; AMYLASEINHTR.
DR SMART; SM00499; AAI; 1.
DR PROSITE; PS00426; CEREAL_1TRYP_1INH; 1.
KW Signal: Alpha-amylase inhibitor.
FT SIGNAL 1 24 POTENTIAL
FT CHAIN 25 171 CMD SUBUNIT OF TETRAMERIC ALPHA-AMYLASE
FT INHIBITOR.
SQ SEQUENCE 171 AA; 18446 MW; A05E5822A3C5480A CRC64;
023982 Length: 171 September 12, 2002 12:10 Type: P Check: 8141 ..
Found using 'consensus' (consens.key)
1 MACKSRSLLLATVMSVFMAAAAAAANDCSPGVAFPVTLGCHRDYVLAQTCVAVLTP
32
61 GSKLEPMWTSALNYPGOPYLAKLYCCQELAKIPQCRCEALRYFMALPVPSQVDPDSTG
121 NVGSGGLMDLPGCPREMGRAVRLVAPGOCNLATINVRCPAVEQPIWI
162

1 match found in sequence:
024000 ; CMD3 PROTEIN PRECURSOR.
(from "consensus_spl.pep")
TOIG of: 024000 check: 7969 from: 1 to: 171
ID 024000 PRELIMINARY; PRT; 171 AA.
AC 024000;
DT 01-JAN-1998 (TREMUREL.05, Created)
DT 01-JAN-1998 (TREMUREL.05, Last sequence update)
DT 01-DEC-2001 (TREMUREL.19, Last annotation update)
DE CMD3 PROTEIN PRECURSOR.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, NK 1558; TISSUE=7 DAYS SEEDLING;
RX MEDLINE=97351567; PubMed=9207849;

RA Grosset J., Alary R., Gautier M.F., Menossi M.,
 RA Martinez-Izquierdo J.A., Joudrier P.;
 RT "Characterization of a barley gene coding for an alpha-amylase
 RT inhibitor subunit (CME protein) and analysis of its promoter in
 RT transgenic tobacco plants and in maize kernels by microprojectile
 RT bombardment.";
 RL Plant Mol. Biol. 34:331-338(1997).
 DR EMBL; U47640; AAB63440.1; -.
 DR HSSP; P01088; 1BEA.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001768; Cereal_tryp_amyl_inh.
 DR Pfam; PF00234; tryp_alpha_amyl_1.
 DR PRINTS; PR00808; AMLASEINHBR.
 DR SMART; SM00499; AAI; 1.
 KW Signal.
 FT SIGNAL. 1 25 POTENTIAL.
 SQ SEQUENCE 171 AA; 18470 MW; A38A5EE2A3C54ECA CRC64;
 024000 Length: 171 September 12, 2002 12:10 Type: P Check: 7969
 Found using 'consensus' (consens.key)

1 -----
 1 match found in sequence:
 049861 : BTI-CME2.1 PROTEIN PRECURSOR.
 (from "consensus-spt.pep")
 TOIG of: 049861 check: 8531 from: 1 to: 148

61 -----
 GSKLPEWMTSAELNYPGOPYLAKLYCCQELAKIPQCRCEALRYFMAIPVPSQPVDPSTG
 32

121 -----
 NVGSGIMLDPGCPREKQAFVRLVAPGOCNLTATHNRYCAVDEPLMI
 162

1 -----
 1 match found in sequence:
 049861 : BTI-CME2.1 PROTEIN PRECURSOR.
 (from "consensus-spt.pep")
 TOIG of: 049861 check: 8531 from: 1 to: 148

AC 049861 PRELIMINARY; PRT; 148 AA.
 DT 01-JUN-1998 (TEMBLrel. 06, Created)
 DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE BTI-CME2.1 PROTEIN PRECURSOR.
 GN ITRI.
 OS Hordeum spontaneum (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_Taxid=77009;
 RN [1]
 RP SROUENCE FROM N.A.
 RC STRAIN-CV. ACCESSION 2; TISSUE-ENDOSPERM;
 RA Rojo J.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ222974; CA11026.1; -.
 DR HSSP; P01087; 1BIU.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001768; Cereal_tryp_amyl_inh.
 DR Pfam; PF00234; tryp_alpha_amyl_1.
 DR PRINTS; PR00808; AMLASEINHBR.
 DR SMART; SM00499; AAI; 1.
 KW Signal.
 FT SIGNAL. 1 24 POTENTIAL.
 FT CHAIN 25 148 BTI-CME2.1 PROTEIN.
 SQ SEQUENCE 148 AA; 16224 MW; C6372793E4285086 CRC64;
 049861 Length: 148 September 12, 2002 12:10 Type: P Check: 8531
 Found using 'consensus' (consens.key)

1 -----
 MAFKYLILSAAYMLAIIATVTSFGDMCAPGDELPHNPLRACRTYVVSQVCHGPRLLT

29 -----
 1 match found in sequence:
 049862 : BTI-CME2.2 PROTEIN PRECURSOR.
 (from "consensus-spt.pep")
 TOIG of: 049862 check: 6829 from: 1 to: 147

61 -----
 SDMKRRCDELSAIPAYCRCEALRIITMGVYWGAFEGAYFDSPPCRERQTSYANL

121 -----
 VTPECNMGTIHGSAYCPELOPGYVVL
 137

1 -----
 1 match found in sequence:
 049863 : BTI-CME4 PROTEIN PRECURSOR.
 (from "consensus-spt.pep")
 TOIG of: 049863 check: 8894 from: 1 to: 148

AC 049863 PRELIMINARY; PRT; 148 AA.
 DT 01-JUN-1998 (TEMBLrel. 06, Created)
 DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE BTI-CME4 PROTEIN PRECURSOR.
 GN ITRI.
 OS Hordeum spontaneum (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_Taxid=77009;
 RN [1]
 RP SROUENCE FROM N.A.
 RC STRAIN-CV. ACCESSION 2; TISSUE-ENDOSPERM;
 RA Rojo J.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ222975; CA11027.1; -.
 DR HSSP; P01087; 1TMO.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR001768; Cereal_tryp_amyl_inh.
 DR Pfam; PF00234; tryp_alpha_amyl_1.
 DR PRINTS; PR00808; AMLASEINHBR.
 DR SMART; SM00499; AAI; 1.
 DR PROSITE; PS00426; CEREAL_TRYP_AMYL_INH. 1.
 KW Signal.
 FT SIGNAL. 1 24 POTENTIAL.
 FT CHAIN 25 147 BTI-CME2.2 PROTEIN.
 SQ SEQUENCE 147 AA; 16199 MW; 888BC989BEBDB9 CRC64;
 049862 Length: 147 September 12, 2002 12:10 Type: P Check: 6829
 Found using 'consensus' (consens.key)

1 -----
 MAFKYLILSAAYMLAIIATVTSFGDMCAPGDELPHNPLRACRTYVVSQVCHGPRLLT
 29

61 -----
 WDMKRRCDELSAIPAYCRCEALRIITMGVYWGAFEGAYFDSPPCRERQTSYANL

121 -----
 NPOECNIMTTHGSPSCPELOPGYVVL
 136

1 -----
 1 match found in sequence:
 049863 : BTI-CME4 PROTEIN PRECURSOR.
 (from "consensus-spt.pep")
 TOIG of: 049863 check: 8894 from: 1 to: 148

AC 049863 PRELIMINARY; PRT; 148 AA.
 DT 01-JUN-1998 (TEMBLrel. 06, Created)
 DT 01-JUN-1998 (TEMBLrel. 06, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE BTI-CME4 PROTEIN PRECURSOR.

GN ITRL.
OS Hordeum spontaneum (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Trilliceae; Hordeum.
OX NCBI_TaxID=77009;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ACCESSION 9; TISSUE-ENDOSPERM;
RA ROYO J.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222976; CAA11028.1; -.
DR HSSP; P01087; ITMO.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal_tryp_aml_1nh.
DR Pfam; PF00234; tryp_alpha_aml_1.
DR PRINTS; PR00808; AMLASEINHBT.
DR SMART; SM00499; AAI; 1.
DR PROSITE; PS00426; CEREAL_TRYP_AMYL_INH; 1.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 148 BTI-CME4.
SQ SEQUENCE 148 AA; 16184 MW; 2515069CEDAA066E CRC64;
049863 Length: 148 September 12, 2002 12:10 Type: P Check: 8894 ..
Found using 'consensus' (consens.key)

1 MAFYQLLISAAMVLAIVATVTSFGDMCAGDALPHNPPLACRTYVVSQICHGPRLLT
29
61 WDMRRCCDELALPAYCRCALRIIMDGIVTWGAFEGAYFKDPNCPRRQTSYANLV
121 TPQECNLMWTIHGSPCELOPGYGVSS
136

1 match found in sequence:
049864 ; BTI-CME1 PRECURSOR.
(from "consensus-spt.pep")
TOIG of: 049864 check: 8460 from: 1 to: 148
ID 049864 PRELIMINARY; PRT; 148 AA.
AC 049864;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BTI-CME1 PRECURSOR.
GN ITRL.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Trilliceae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. ALBACETE; TISSUE-ENDOSPERM;
RA ROYO J.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222977; CAA11029.1; -.
DR HSSP; P01087; ITMO.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal_tryp_aml_1nh.
DR Pfam; PF00234; tryp_alpha_aml_1.
DR PRINTS; PR00808; AMLASEINHBT.
DR SMART; SM00499; AAI; 1.
DR PROSITE; PS00426; CEREAL_TRYP_AMYL_INH; 1.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 148 BTI-CME1.
SQ

SQ SEQUENCE 148 AA; 16365 MW; ABFFD7CD12F07ADI CRC64;
049864 Length: 148 September 12, 2002 12:10 Type: P Check: 8460 ..
Found using 'consensus' (consens.key)

1 MAFYQLLISAAMVLAIVATVTSFGDMCAGDALPHNPPLACRTYVVSQICHGPRLLT
29
61 SDMRRCCDELALPAYCRCEALRIIMGVVWGAFAEGAYFKDPNCPRRQTSYANLV
121 VTQECNLMWTIHGSPCELOPGYGVSS
137

1 match found in sequence:
049865 ; BTI-CME3.1 PROTEIN PRECURSOR.
(from "consensus-spt.pep")
TOIG of: 049865 check: 1427 from: 1 to: 149
ID 049865 PRELIMINARY; PRT; 149 AA.
AC 049865;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE BTI-CME3.1 PROTEIN PRECURSOR.
GN ITRL.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Trilliceae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. VALTICKY; TISSUE-ENDOSPERM;
RA ROYO J.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222978; CAA11030.1; -.
DR HSSP; P01087; ITMO.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal_tryp_aml_1nh.
DR Pfam; PF00234; tryp_alpha_aml_1.
DR PRINTS; PR00808; AMLASEINHBT.
DR SMART; SM00499; AAI; 1.
DR PROSITE; PS00426; CEREAL_TRYP_AMYL_INH; 1.
KW Signal.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 149 BTI-CME3.1 PROTEIN.
SQ SEQUENCE 149 AA; 16290 MW; C61A890CEB9E536F CRC64;
049865 Length: 149 September 12, 2002 12:10 Type: P Check: 1427 ..
Found using 'consensus' (consens.key)
1 MAFYQLLISAAMVLAIVATVTSFGDMCAGDALPHNPPLACRTYVVSQICHGPRLLT
29
61 SDMRRCCDELALPAYCRCEALRIIMGVVWGAFAEGAYFKDPNCPRRQTSYANLV
121 VTQECNLMWTIHGSPCELOPGYGVSS
137

1 match found in sequence:
049867 ; BTI-CME2.2 PROTEIN PRECURSOR.
(from "consensus-spt.pep")
TOIG of: 049867 check: 7816 from: 1 to: 147

```

ID 049867 PRELIMINARY; PRT; 147 AA.
AC 049867;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE BTI-CME2.2 PROTEIN PRECURSOR.
GN TRYI.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_Taxid=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. HATIF DE GRIGNON; TISSUE-ENDOSPERM;
RA ROYO J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ223458; CAA1373.1; -.
DR HSSP; P01087; 1TMO.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal tryp_aml_1nh.
DR Pfam; PF00234; tryp_alpha_aml1; 1.
DR PRINTS; PR00808; AMLASEINH8TR.
DR SMART; SM00499; AAI; 1.
DR PROSITE; PS00426; CEREAL_TRYP_AMYL_INH; 1.
KW Signal.
FT CHAIN 1 24 POTENTIAL.
FT CHAIN 25 147 BTI-CME2.2 PROTEIN.
SQ SEQUENCE 147 AA; 16171 MW; 87B69300AF4DDEB9 CRC64;

049867 Length: 147 September 12, 2002 12:10 Type: P Check: 7816
Found using 'consensus' (consens.key)

```

```

1 MAKYQLLLSAAVLAATLATVTSFGDMCAPGADLPANPLACTYVVSQICHWGPRIST
29
-----
61 WDMKRRCCDELSAIPAYCRCALRIINDGVTWQGVGAYFKDMPNSPRVWQTSYANLV
-----
121 NPOECNLMTIHGSPSCPELOPGYEVL
136
-----
1 match found in sequence:
p80208; NAPIN BNITI (1.7S SEED STORAGE PROTEIN).
(from "consensus-spt.pep")
TOIG of: p80208 check: 2142 from: 1 to: 125

ID p80208 PRELIMINARY; PRT; 125 AA.
AC p80208;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE NAPIN BNITI (1.7S SEED STORAGE PROTEIN).
OS Brassica napus (Rape).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Brassica.
OX NCBI_Taxid=3708;
RN [1]
RP SEQUENCE.
RC TISSUE=SEED.
RA MIDLIN=97390373; PubMed=9249277;
RA Monaldi R.I., Gonzalez de la Pena M.A., Lopez-Otin C., Flanador A.,
RA Fernandez C., Villalba M., Rodriguez R.;
RT "Detection, isolation and complete amino acid sequence of an
RT allergenic protein from rapeseed flour.";
RL Clin. Exp. Allergy 27:833-841(1997).
CC -1- FUNCTION: THE SMALL, BASIC, WATER-SOLUBLE NAPINS ARE ONE OF THE

```

```

CC TWO MAJOR KINDS OF STORAGE PROTEINS SYNTHESIZED IN THE SEED DURING
CC ITS MATURATION.
CC -1- SUBUNIT: THE MATURE PROTEIN CONSISTS OF A SMALL AND A LARGE CHAIN
CC LINKED BY TWO DISULFIDE BONDS.
CC -1- TISSUE SPECIFICITY: COTYLEDONS AND THE AXIS.
CC -1- SIMILARITY: BELONGS TO THE 2S SEED STORAGE ALBUMINS FAMILY.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal tryp_aml_1nh.
DR InterPro; IPR000617; Napin.
DR Pfam; PF00234; tryp_alpha_aml1; 1.
DR PRINTS; PR00496; NAPIN.
DR PRODOM; PD002498; Napin; 1.
DR SMART; SM00499; AAI; 1.
KW Seed storage protein; Multigene family.
FT CHAIN 1 37 SMALL CHAIN.
FT CHAIN 38 125 LARGE CHAIN.
SQ SEQUENCE 125 AA; 14035 MW; C4042A1BF07099B8 CRC64;

p80208 Length: 125 September 12, 2002 12:10 Type: P Check: 2142
Found using 'consensus' (consens.key)

```

```

1 SAGFRIPKCRKEHQAGHRLACQWHLKQAMQSGSGPQGRPRPLQCCCNELHQEEP
10
-----
61 LCVCPTLKGSRAVKQGVROQGGQQLQGVISRTVQTATHLPKVCNIPQVSCPKT
115
-----
121 MPGPS
-----

```

```

1 match found in sequence:
p93198; ALBUMIN SEED STORAGE PROTEIN (FRAGMENT).
(from "consensus-spt.pep")
TOIG of: p93198 check: 7995 from: 1 to: 139

```

```

ID p93198 PRELIMINARY; PRT; 139 AA.
AC p93198;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ALBUMIN SEED STORAGE PROTEIN (FRAGMENT).
OS Juglans regia (English walnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Juglandaceae; Juglans.
OX NCBI_Taxid=51240;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SUNLAND;
RA Teuber S.S., Dandekar A.M., Peterson W.R.;
RT "Juglans regia 2S albumin seed storage protein precursor.";
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U66866; AABA1308.1; -.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal tryp_aml_1nh.
DR Pfam; PF00234; tryp_alpha_aml1; 1.
DR PRINTS; PR00211; GLUTELIN.
DR SMART; SM00499; AAI; 1.
FT NON TER 1 1
SQ SEQUENCE 139 AA; 16373 MW; 02D0E55E67164F23 CRC64;

p93198 Length: 139 September 12, 2002 12:10 Type: P Check: 7995
Found using 'consensus' (consens.key)

```

```

1 ALLVALLFVANAARFTTTTWEIDEDIDNPRRGECEQIOPRQONLHNCQYLRQOS
39
-----
61 RSGGYDEDNORHFRCCOOLSDMECCCEGLROYVRRQOQOGLRGEEMEVQASD
-----

```

```

121 -----|
      LPNECGISSQRCIEIRRSWF
      132

-----|
1 match found in sequence:
P93414 ; 26 KDA GLOBULIN.
      (from "consensus_spt.pep")
TOIG of: P93414 check: 1219 from: 1 to: 186

ID P93414 PRELIMINARY; PRT; 186 AA.
AC P93414;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE 26 KDA GLOBULIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Erihartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOSPERM;
RX MEDLINE=96235139; PubMed=8666249;
RA Nakase M., Hotta H., Adachi T., Aoki N., Nakamura R., Masumura T.,
RA Tanaka K., Matsuda T.;
RT "Cloning of the rice seed alpha-globulin-encoding gene: sequence
RT similarity of the 5'-flanking region to those of the genes encoding
RT wheat high molecular-weight glutenin and barley D hordein.";
RL Gene 170:223-226(1996).
DR EMBL, D50643; BAA09308.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; tryp_alpha_amy1_1.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 186 AA; 21055 MW; AE2B8F1107C8BC94 CRC64;

P93414 Length: 186 September 12, 2002 12:10 Type: P Check: 1219 ..
Found using 'consensus' (consens.key)

1 -----|
      MASKVFFFAALMAMVAISGAQLSESEMRPRRCOREVDSPIDACRQVLDRLTGRE
      36
61 -----|
      RPOPFRRPAGLGLRMCCQOLQDYRSRCRAIRRMVRSYEESMPMLEQGWSSSSEY
121 -----|
      YGEGSSSEQGYGEGSSEGYEGEQQPGKTRVRLTRANQYAAQLPSCMRVPEQCSI
      178
181 FAAGOY

-----|
1 match found in sequence:
q39787 ; 2S ALBUMIN STORAGE PROTEIN PRECURSOR.
      (from "consensus_spt.pep")
TOIG of: q39787 check: 6563 from: 1 to: 139

ID Q39787 PRELIMINARY; PRT; 139 AA.
AC Q39787;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 2S ALBUMIN STORAGE PROTEIN PRECURSOR.
GN Mat5-D.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Spentatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

1 -----|
      MAKLAVALTALILFLANASITSVTVESEENRDCSEQIRKQALHKQKYEEMWGE
      36
61 -----|
      GSDNIAGYIDSCCOLEKMDQCRCQGLRHATMOOMQOMQOMKQREIMQVTKNI
121 -----|
      MSECMEPRCDTPERSLI
      131

-----|
1 match found in sequence:
q39795 ; 2S ALBUMIN STORAGE PROTEIN PRECURSOR.
      (from "consensus_spt.pep")
TOIG of: q39795 check: 6011 from: 1 to: 139

ID Q39795 PRELIMINARY; PRT; 139 AA.
AC Q39795;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 2S ALBUMIN STORAGE PROTEIN PRECURSOR.
GN Mat5-A.
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC Eucosids II; Malvales; Malvaceae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COTYLEDON;
RX Galau G.A., Wang H.Y.-C., Hughes D.W.;
RT "Cotton Mat5-A (C164) gene and Mat5-D cDNAs encoding methionine-rich
RT 2S albumin storage proteins.";
RL Submitted (JAN-1992) to the EMBL/Genbank/DBJ databases.
DR EMBL; M86213; AAA33066.1; -.
DR InterPro; IPR003612; AAI.
DR Pfam; PF00234; tryp_alpha_amy1_1.
DR SMART; SM00499; AAI; 1.
DR SIGNAL.
KW SIGNAL.
FT CHAIN 1 20 POTENTIAL.
FT CHAIN 33 59 2S ALBUMIN STORAGE PROTEIN.
FT CHAIN 64 139 2S ALBUMIN STORAGE PROTEIN.
SQ SEQUENCE 139 AA; 15700 MW; 02ACE24FE9CE9F90 CRC64;

Q39795 Length: 139 September 12, 2002 12:10 Type: P Check: 6011 ..
Found using 'consensus' (consens.key)

```

```

1  -----|-----
1  MAXLAVALATLALLFLANASITSVTVESENRDSCSEQOIKQKHLKCKQYMEELGGE
36
61  GSDNIAGGYIDSCCOLEKMDTOCRGCRGLRHATWQMOOMOGOMGSKOMREIMOKVTKI
-----|-----
121  MSECMEPRGCDTPSRSLI
131
1 match found in sequence:
q41540 : CM 17 PROTEIN.
(from "consensus_spt.pep")
TOIG of: q41540 check: 6711 from: 1 to: 143

ID 041540 PRELIMINARY; PRT; 143 AA.
AC 041540;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE CM 17 PROTEIN.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_Taxid=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STEFANIN-CV, TRMCALEN;
RA Lullien V., Alary R., Guirao A., Joudrier P., Gautier M.F.;
RT "Isolation and nucleotide sequence of a cDNA-clone encoding the bread
RT wheat (Triticum aestivum L.) CM17 protein.";
RL Plant Mol. Biol. 170:1081-1082(1991).
DR EMBL; X59791; CAA42453.1; -
DR HSSP; P01088; IBEA.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_amy1_inh.
DR Pfam; PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00808; AMLASEINHTR.
DR SMART; SM00499; AAI; 1.
DR PROSITE; PS00426; CEREAL_TRYP_AMYL_INH; 1.
FT CHAIN 25 143 CM 17 PROTEIN.
SQ SEQUENCE 143 AA; 15989 MW; 4F1A426EADDF09B CRC64;

041540 Length: 143 September 12, 2002 12:10 Type: P Check: 6711 ..
Found using 'consensus' (consens.key)

1  -----|-----
1  MASKSNYLLFTALLVFIPAAVAAGVNDGCPWMTSTLITPLPSCRNVEBQACNIEMPGP
30
61  PYLAKQCCBCEQLANIPOOQRQALRYFMGPKSRPDQSGIMELPGCPREVQWNPVILVTP
-----|-----
121  GYCNLTIVHNPYICLGMESQMS
134
1 match found in sequence:
q946v1 : HAGEMAN FACTOR INHIBITOR.
(from "consensus_spt.pep")
TOIG of: q946v1 check: 3458 from: 1 to: 185

ID 0946v1 PRELIMINARY; PRT; 185 AA.
AC 0946v1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)

```

```

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HAGEMAN FACTOR INHIBITOR.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoidae; Andropogoneae; Zea.
OX NCBI_Taxid=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-ENDOSPERM;
RX MEDLINE=21480069; Pubmed=11595803;
RA Woo Y.M., Hu D.W., Larkins B.A., Jung R.;
RT "Genomics analysis of genes expressed in maize endosperm identifies
RT novel seed proteins and clarifies patterns of zein gene expression.";
RL Plant Cell 13:2297-2317(2001).
DR EMBL; AF371280; AAL16995.1; -
SQ SEQUENCE 185 AA; 19626 MW; E5762061D611FE35 CRC64;

0946v1 Length: 185 September 12, 2002 12:10 Type: P Check: 3458 ..
Found using 'consensus' (consens.key)

...
14  NANCCKRSIERSTGELMASSSSSHRRLLILAAVLLSVLAASASAGTSCVPGMAIPHN
64
74  PLPSCRNVYSRTGIGIRPLPWPBLKRCRCRELADIPAYRCRTALSLMGCALIPGPDQ
-----|-----
134  LEGRLIEDPGCPREVQGFATLVTEACMLATISGVAECPEWILGGTMSPK
173
1 match found in sequence:
q9aqx9 : PUROINDOLINE B.
(from "consensus_spt.pep")
TOIG of: q9aqx9 check: 6622 from: 1 to: 148

ID 09AQX9 PRELIMINARY; PRT; 148 AA.
AC 09AQX9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE PUROINDOLINE B.
GN PINB.
OS Triticum urartu, and
OS Triticum monococcum (Einkorn wheat) (Small spelt).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_Taxid=4572, 4568;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=T.urartu, and T.monococcum;
RA Lillien M., Simeone M.C., Morris C.F.;
RT "Analysis of puroindoline a and b sequences from Triticum aestivum cv.
RT 'Penawawa' and diploid wheat ancestors.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ302102; CAC33800.1; -
DR EMBL; AJ302102; CAC33800.1; -
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_amy1_inh.
DR Pfam; PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00808; AMLASEINHTR.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 148 AA; 16682 MW; B4ECA6C4C9ACE059 CRC64;

09AQX9 Length: 148 September 12, 2002 12:10 Type: P Check: 6622 ..
Found using 'consensus' (consens.key)

```

```

1      MKTLFLALLALVASTTFAQYSEVGVNENYAGAGSGSQOCPLEPRKLSCKDYMERCF
-----|-----
61      MKDPPVPTKMKWKGGEHEVERKCCOOLSOIAPQCRDLSIRGMIOKLGFGFIMRGDV
-----|-----
121     FKQIQRAQSLPSKCNMGADCKLPSCGYW
-----|-----
140

1 match found in sequence:
q9aup8 : 2S ALBUMIN.
(from "consensus_spt.pep")
TOIG of: q9aup8 check: 5064 from: 1 to: 153

ID 09AUD1 PRELIMINARY; PRT; 153 AA.
AC 09AUD1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE 2S ALBUMIN.
OS Sesamum indicum (Oriental sesame) (Singelly).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Lamiales; Pedaliaceae; Sesamum.
OX NCBI_TaxID=4182;
RN [1]
RP SEQUENCE FROM N.A.
RA Tai S.S.K., Tzen J.T.C.;
RT "Molecular cloning of three storage proteins in sesame.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF240005; AAK15088.1; -.
DR HSSP; P01085; 1HSS.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_aml1_inh.
DR InterPro: IPR000480; Glutelin.
DR InterPro: IPR000617; Napin.
DR Pfam; PF00234; tryp_alpha_aml1; 1.
DR PRINTS; PR00211; GLUTELIN.
DR PRINTS; PR00496; NAPIN.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 153 AA; 17504 MW; 6A8F9117DAE81568 CRC64;

09AUD1 Length: 153 September 12, 2002 12:10 Type: P Check: 5064
Found using 'consensus' (consens.key)

1      MAKTLALAVILVAMVALASATTTTATTTAIDDEANQSQOCCROOLQGRSCORYL
-----|-----
44

61      SQGRSPYGGGEDEVLEMTSGNQSQESQSLRDCCQLRWVDECRCEALRQAVRQOQGGY
-----|-----

121     QEGSQGVYQARALPRRCNMRPOCCOFRVLFV
-----|-----
146

1 match found in sequence:
q9aup8 : PUROINDOLINE B.
(from "consensus_spt.pep")
TOIG of: q9aup8 check: 6382 from: 1 to: 148

ID 09AVP8 PRELIMINARY; PRT; 148 AA.
AC 09AVP8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PUROINDOLINE B.

```

```

GN PINB.
OS Triticum urartu.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4572;
RN [1]
RP SEQUENCE FROM N.A.
RA Lillomo M., Simeone M.C., Morris C.F.;
RT "Analysis of puroindoline a and b sequences from Triticum aestivum cv.
RT 'Penawawa' and diploid wheat ancestors.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ302103; CAC33799.1; -.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_aml1_inh.
DR Pfam; PF00234; tryp_alpha_aml1; 1.
DR PRINTS; PR00808; AMTASEINHTR.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 148 AA; 16639 MW; B4ECA6D2041A2DP9 CRC64;

09AVP8 Length: 148 September 12, 2002 12:10 Type: P Check: 6382
Found using 'consensus' (consens.key)

1      MKTLFLALLALVASTTFAQYSEVGVNENYAGAGSGSQOCPLEPRKLSCKDYMERCF
-----|-----
40

61      MKDPPVPTKMKWKGGEHEVERKCCOOLSOIAPQCLDLSIRGMIOKLGFGFIMRGDV
-----|-----
140

121     FKQIQRAQSLPSKCNMGADCKLPSCGYW
-----|-----
140

1 match found in sequence:
q9aup9 : PUROINDOLINE B.
(from "consensus_spt.pep")
TOIG of: q9aup9 check: 6496 from: 1 to: 148

ID 09AVP9 PRELIMINARY; PRT; 148 AA.
AC 09AVP9;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PUROINDOLINE B.
GN PINB.
OS Triticum monococcum (Einkorn wheat) (Small spelt).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Triticum.
OX NCBI_TaxID=4568;
RN [1]
RP SEQUENCE FROM N.A.
RA Lillomo M., Simeone M.C., Morris C.F.;
RT "Analysis of puroindoline a and b sequences from Triticum aestivum cv.
RT 'Penawawa' and diploid wheat ancestors.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ302101; CAC33795.1; -.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_aml1_inh.
DR Pfam; PF00234; tryp_alpha_aml1; 1.
DR PRINTS; PR00808; AMTASEINHTR.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 148 AA; 16654 MW; B4ECA6CADFDD8759 CRC64;

09AVP9 Length: 148 September 12, 2002 12:10 Type: P Check: 6496
Found using 'consensus' (consens.key)

1      MKTLFLALLALVASTTFAQYSEVGVNENYAGAGSGSQOCPLEPRKLSCKDYMERCF
-----|-----
40

```

```

61 MKDPVTPWKMKGGCEHVEKCKCKOLSOIAPQCRCDLSLRVIOGKLGFGIMRGVGA
-----|-----
121 FKQIORAOSLPSKCMGADCKPSPGYW
140
-----|-----
1 match found in sequence:
  q9awg9 ; PUROINDOLINE B.
  (from "consensus-spt.pep")
  TOIG of: q9awg9 check: 6685 from: 1 to: 148

ID O9AWG9 PRELIMINARY; PRT; 148 AA.
AC O9AWG9;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)
DE PUROINDOLINE B.
GN PINB.
OS Aegilops speltoides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Aegilops.
NCBI_TaxID=4573;
[1]
RP SEQUENCE FROM N.A.
RA Lillmo M., Simeone M.C., Morris C.F.;
RT "Analysis of puroindoline a and b sequences from Triticum aestivum cv.
RT 'Penawawa' and diploid wheat ancestors.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ302105; CAC33505.1; -
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_amy1_inh.
DR Pfam: PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00808; AMLASEINHBR.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 148 AA; 16703 MW; 0151E2028DAD6AA9 CRC64;

O9AWG9 Length: 148 September 12, 2002 12:10 Type: P Check: 6685
Found using 'consensus' (consens.key)

1 MKSLFLGILALVASTTFAQYSEVGWYNEVGSGSQCCPOERPKLGSKDYLERCFT
40
-----|-----
61 MKDPVTPWKMKGGCEHVEKCKCKOLSOIAPQCRCDLSLRVIOGKLGFGIMRGV
-----|-----
121 FKQIORAOSLPSKCMGADCKPSPGYW
140
-----|-----
1 match found in sequence:
  q9awh0 ; PUROINDOLINE B.
  (from "consensus-spt.pep")
  TOIG of: q9awh0 check: 3610 from: 1 to: 147

ID O9AWH0 PRELIMINARY; PRT; 147 AA.
AC O9AWH0;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)
DE PUROINDOLINE B.
GN PINB.
OS Aegilops speltoides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Aegilops.

```

```

OX NCBI_TaxID=4573;
RN [1]
RP SEQUENCE FROM N.A.
RA Lillmo M., Simeone M.C., Morris C.F.;
RT "Analysis of puroindoline a and b sequences from Triticum aestivum cv.
RT 'Penawawa' and diploid wheat ancestors.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ302105; CAC33505.1; -
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_amy1_inh.
DR Pfam: PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00808; AMLASEINHBR.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 147 AA; 16638 MW; FFF7131EBF523BEF8 CRC64;

O9AWH0 Length: 147 September 12, 2002 12:10 Type: P Check: 3610
Found using 'consensus' (consens.key)

1 MKSLFLGILALVASTTFAQYSEVGWYNEVGSGSQCCPOERPKLGSKDYLERCFTM
39
-----|-----
61 KDPVTPWKMKGGCEHVEKCKCKOLSOIAPQCRCDLSLRVIOGKLGFGIMRGV
-----|-----
121 KOIORAOSLPSKCMGADCKFPPIGYW
139
-----|-----
1 match found in sequence:
  q9fs19 ; HORDINDOLINE-B.
  (from "consensus-spt.pep")
  TOIG of: q9fs19 check: 2569 from: 1 to: 147

ID O9FS19 PRELIMINARY; PRT; 147 AA.
AC O9FS19;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, last sequence update)
DE HORDINDOLINE-B.
GN H01-B1.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticaceae; Hordeum.
NCBI_TaxID=4513;
[1]
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-ETOLATED LEAVES;
RA Rouster J., Darlington H.F., Shewry P.R., Simpson D.;
RT "Hordeolindolines-B.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ276143; CAB92291.2; -
DR HSSP; P01087; 1B10.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_amy1_inh.
DR Pfam: PF00234; tryp_alpha_amy1; 1.
DR PRINTS; PR00808; AMLASEINHBR.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 147 AA; 16122 MW; 57871761245CDD40 CRC64;

O9FS19 Length: 147 September 12, 2002 12:10 Type: P Check: 2569
Found using 'consensus' (consens.key)

1 MKTLFLALIALVASTTFAQYSEVGWYNEVGSGSQCCPOERPKLGSKDYLERCFTM
39
-----|-----
61 KDPVTPWKMKGGCEHVEKCKCKOLSOIAPQCRCDLSLRVIOGKLGFGIMRGV
-----|-----

```

121 KOIOIRAQILPSKCMGADCKRPSGYW
139

1 match found in sequence:
q9le34 : HORDOIINDOLINE (FRAGMENT).

(from "consensus_spt.pep")
TOIG of: q9le34 check: 8354 from: 1 to: 141

ID Q9LE34 PRELIMINARY; PRT; 141 AA.
AC Q9LE34;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HORDOIINDOLINE (FRAGMENT).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, TRIUMPH, AND CV, CHARLOT, TISSUE-ENDOSPERM;
RA Rousier J., Darlington H.F., Shewry P.R., Simpson D.;
RT "Hordeindolines";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277550; CAB90638.1; -
DR EMBL; AJ277549; CAB90637.1; -
DR HSSP; P01087; 1B1U.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal_1ryp-amy1_inh.
DR Pfam; PF00234; 1ryp_alpha_amy1; 1.
DR PRINTS; PR00808; AMLASEINHTR.
DR SMART; SM00499; AAI; 1.
FT NON_TER 1
FT CHAIN 31 >141 HORDOIINDOLINE.
FT NON_TER 141 141
SQ SEQUENCE 141 AA; 15164 MW; 24BA20D710A5B283 CRC64;

Q9LE34 length: 141 September 12, 2002 12:10 Type: P Check: 8354
Found using 'consensus' (consens.key)

1 GSKMTFLALALVASTSAOYVGGYNDVGGGGGSOCCPQERPNLGSCKDYMERCF
41

61 TMADPVTWPTKWKKGCEHEVREKCCQQLSQIAPHCRDAIRGVIOGKLGIFGIGGA

121 VFKIOIRAQILPSKCMNGVDC
141

1 match found in sequence:

q9le8 : HORDOIINDOLINE-B.
(from "consensus_spt.pep")

TOIG of: q9le8 check: 3041 from: 1 to: 147

ID Q9LE8 PRELIMINARY; PRT; 147 AA.
AC Q9LE8;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HORDOIINDOLINE-B.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ETOLATED LEAVES;
RA Rousier J., Darlington H.F., Shewry P.R., Simpson D.;
RT "Hordeindolines";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ276143; CAB92292.1; -
DR HSSP; P01087; 1B1U.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal_1ryp-amy1_inh.
DR Pfam; PF00234; 1ryp_alpha_amy1; 1.
DR PRINTS; PR00808; AMLASEINHTR.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 147 AA; 16078 MW; C106BDB5A7CFD22 CRC64;

Q9LE8 length: 147 September 12, 2002 12:10 Type: P Check: 3041
Found using 'consensus' (consens.key)

1 MKTLFLALALVASTSAOYVGGYNDVGGGGGSOCCPQERPNLGSCKDYMERCFM
39

61 KDFPVTWPTKWKKGCEHEVREKCCQQLSQIAPHCRDAIRGVIOGKLGIFGIGGAVF

121 KOIOIRAQILPSKCMNGVDCRPSGYW
139

1 match found in sequence:
q9lrc2 : 2S ALBUMIN.

(from "consensus_spt.pep")
TOIG of: q9lrc2 check: 1726 from: 1 to: 146

ID Q9LRC2 PRELIMINARY; PRT; 146 AA.
AC Q9LRC2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE 2S ALBUMIN.
OS Bertholletia excelsa (Brazil nut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; Ericales; Lecythidaceae; Bertholletia.
OX NCBI_TaxID=3645;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamauchi D.;
RT "Brazil nut 2S albumin was synthesized in a transgenic French bean seed with a promoter of the gene for canavallin, 7S globulin from
RT Canavalia gladiata";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB044391; BAA96554.1; -
DR HSSP; P01087; 1B1U.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001768; Cereal_1ryp-amy1_inh.
DR InterPro; IPR000617; Napin.
DR Pfam; PF00234; 1ryp_alpha_amy1; 1.
DR PRINTS; PR00808; AMLASEINHTR.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 146 AA; 16910 MW; 4A69A196E6EC7096 CRC64;

Q9LRC2 length: 146 September 12, 2002 12:10 Type: P Check: 1726
Found using 'consensus' (consens.key)

1 MAKISVAALALVMAIGHATAFRATVTTVEEENOECEKREQOMROOMLSHCMMYRROO

40

61 MESPPTWPRGMEPRMSSECEQLCEMDSCREGLRMWMMQKQEMQPRGQMRM

121 RLAENIPRCNLSPRCMGSGIAGF
137

1 match found in sequence:

q9m3v3 ; HORDOINDOLINE (FRAGMENT).
(from "consensus-spt.pep")
TOIG of: q9m3v3 check: 8764 from: 1 to: 141

ID Q9M3V3 PRELIMINARY; PRT; 141 AA.
AC Q9M3V3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE HORDOINDOLINE (FRAGMENT).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SUNDANCE; TISSUE-ENDOSPERM;
RA Rousier J., Darlington H.F., Shewry P.R., Simpson D.;
RT "Hordeindolines";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277553; CAB90641.1; -
DR HSSP: P01087; 1B1U.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_aml1_inh.
DR InterPro: IPR000480; Glutelin.
DR Pfam: PF00234; tryp_alpha_aml1; 1.
DR PRINTS: PR00808; AMLASEINHBR.
DR PRINTS: PR00211; GLUTELIN.
DR SMART: SM00499; AAI; 1.
FT NON_TER 1
FT CHAIN 31 >141 HORDOINDOLINE.
FT NON_TER 141 141
SQ SEQUENCE 141 AA; 15263 MW; 62CAB45D0B6AECB5 CRC64;
O9M3V3 Length: 141 September 12, 2002 12:10 Type: P Check: 8764
Found using 'consensus' (consens.key)

1 SMKTLFLALLALVASTTSQAQSVGGGYNVGGGSGSQCPQERPNLGSCKDYMERCF
40

61 MKDFPTWPTKMKWGCEHEVRKCCQQLSQIAPHRCDAIRGVIOGKLGIFGIGGAV

121 FKQIQRAQILPSCKNMGVDCR
140

1 match found in sequence:
q9m3v4 ; HORDOINDOLINE (FRAGMENT).
(from "consensus-spt.pep")

TOIG of: q9m3v4 check: 578 from: 1 to: 142

ID Q9M3V4 PRELIMINARY; PRT; 142 AA.
AC Q9M3V4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE HORDOINDOLINE (FRAGMENT).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. HART; TISSUE-ENDOSPERM;
RA Rousier J., Darlington H.F., Shewry P.R., Simpson D.;
RT "Hordeindolines";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277552; CAB90640.1; -
DR HSSP: P01087; 1B1U.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_aml1_inh.
DR InterPro: IPR000480; Glutelin.
DR Pfam: PF00234; tryp_alpha_aml1; 1.
DR PRINTS: PR00808; AMLASEINHBR.
DR PRINTS: PR00211; GLUTELIN.
FT NON_TER 1
FT CHAIN 31 >142 HORDOINDOLINE.
FT NON_TER 142 142
SQ SEQUENCE 142 AA; 15331 MW; 9D94BA20D410A4FB CRC64;
O9M3V4 Length: 142 September 12, 2002 12:10 Type: P Check: 578
Found using 'consensus' (consens.key)

1 GSKMTFLALLALVASTTSQAQIPVGGGYNVGGGSGSQCPQERPNLGSCKDYMERCF
41

61 TMKDFPTWPTKMKWGCEHEVRKCCQQLSQIAPHRCDAIRGVIOGKLGIFGIGGAV

121 VFQIQRAQILPSCKNMGVDCR
141

1 match found in sequence:
q9m3v5 ; HORDOINDOLINE (FRAGMENT).
(from "consensus-spt.pep")
TOIG of: q9m3v5 check: 793 from: 1 to: 142

ID Q9M3V5 PRELIMINARY; PRT; 142 AA.
AC Q9M3V5;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HORDOINDOLINE (FRAGMENT).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DERRADO; TISSUE-ENDOSPERM;
RA Rousier J., Darlington H.F., Shewry P.R., Simpson D.;
RT "Hordeindolines";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ277551; CAB90639.1; -
DR HSSP: P01087; 1B1U.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; Cereal_tryp_aml1_inh.
DR InterPro: IPR000480; Glutelin.
DR Pfam: PF00234; tryp_alpha_aml1; 1.
DR PRINTS: PR00808; AMLASEINHBR.
DR PRINTS: PR00211; GLUTELIN.
DR SMART: SM00499; AAI; 1.

FT NON_TER 1 1
 FT CHAIN 31 >142 HORDOINDOLINE.
 FT NON_TER 142
 SO SEQUENCE 142 AA; 15307 MW; AE82D6F0D710A5B2 CRC64;

Q9M3V5 Length: 142 September 12, 2002 12:10 Type: P Check: 793
 Found using 'consensus' (consens.key)

1 GSMKTLFLALALVASTSAQYVGGGYNVGGGSGSQCPQERPNLGSCDYMERCF
 41

61 TMKDFPTWPTKMKKGCEHEVREKCCQQLSQIAPHRCDAIRGVIOGKLGIGIGGGA

121 VFKQVQRAQLIPSKCNMGVDCR
 141

1 match found in sequence:
 q9m3v6 ; HORDOINDOLINE (FRAGMENT).
 (from "consensus-spt.pep")
 TOIG of: q9m3v6 check: 650 from: 1 to: 142

ID Q9M3V6 PRELIMINARY; PRT; 142 AA.
 AC Q9M3V6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HORDOINDOLINE (FRAGMENT).
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_Taxid=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, BRITANNIA; TISSUE=ENDOSPERM;
 RA Kousler J., Darlington H.F., Shewry P.R., Simpson D.;
 RT "Hordeindolines."
 RL EMBL: AJ277548; CAB90636.1; -
 DR EMBL: AJ277548; CAB90636.1; -
 DR HSP: P01087; IRIU
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereal_1tryp_aml_1nh.
 DR InterPro: IPR000480; Glutelin.
 DR Pfam: PF00234; tryp_alpha_aml1; 1.
 DR PRINTS: PR00808; AMLASEINHTR.
 DR PRINTS: PR00211; GLUTELIN.
 DR SMART: SM00499; AAI; 1.
 FT NON_TER 1
 FT CHAIN 31 >142 HORDOINDOLINE.
 FT NON_TER 142
 SO SEQUENCE 142 AA; 15321 MW; AE94BA20D710A5B2 CRC64;

Q9M3V6 Length: 142 September 12, 2002 12:10 Type: P Check: 650
 Found using 'consensus' (consens.key)

1 GSMKTLFLALALVASTSAQYVGGGYNVGGGSGSQCPQERPNLGSCDYMERCF
 41

61 TMKDFPTWPTKMKKGCEHEVREKCCQQLSQIAPHRCDAIRGVIOGKLGIGIGGGA

121 VFKQVQRAQLIPSKCNMGVDCR
 141

1 match found in sequence:
 q9m4d9 ; PUROINDOLINE-B PRECURSOR (PUROINDOLINE B).
 (from "consensus-spt.pep")
 TOIG of: q9m4d9 check: 6986 from: 1 to: 148

ID Q9M4D9 PRELIMINARY; PRT; 148 AA.
 AC Q9M4D9;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUROINDOLINE-B PRECURSOR (PUROINDOLINE B).
 OS MONOPYRUB OR PINB.
 OS Triticum monococcum (Einkorn wheat) (Small spelt), and
 OS Aegilops tauschii (Aegilops squarrosa).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 OC NCBI_Taxid=4568; 37682;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=T.monococcum; TISSUE=LEAVES;
 RA Gautier M.F., Cosson P., Guirao A., Alary R., Joudrier P.;
 RT "Puroindoline genes are highly conserved in diploid ancestor wheats
 and related species but absent in tetraploid Triticum species."
 RL Plant Sci. 153:81-91(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=A.tauschii;
 RA Lillamo M., Simeone M.C., Morris C.F.;
 RT "Analysis of puroindoline a and b sequences from Triticum aestivum cv.
 'Penawawa' and diploid wheat ancestors."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ249934; CAB89542.1; -
 DR EMBL: AJ302108; CAC33509.1; -
 DR EMBL: AJ302107; CAC33509.1; -
 DR InterPro: IPR001768; AAI.
 DR InterPro: IPR001768; Cereal_1tryp_aml_1nh.
 DR Pfam: PF00234; tryp_alpha_aml1; 1.
 DR PRINTS: PR00808; AMLASEINHTR.
 DR SMART: SM00499; AAI; 1.
 FT SIGNAL
 FT CHAIN 1 19 POTENTIAL.
 FT NON_TER 30 148 PUROINDOLINE-B.
 SO SEQUENCE 148 AA; 16738 MW; 3B7B02008591A1D CRC64;

Q9M4D9 Length: 148 September 12, 2002 12:10 Type: P Check: 6986
 Found using 'consensus' (consens.key)

1 MKTLFLALALVASTTFAQYSEVGMVNEVAGAGSSQCPLEPRKLSCKDYMGCMFT
 40

61 MKDFPTWPTKMKKGCEHEVREKCCQQLSQIAPHRCDSIRGMIOGKLGIFGIWRDV

121 FFKIQRAQLIPSKCNMGADCKFPFGIYW
 140

1 match found in sequence:
 q9m4d4 ; HORDOINDOLINE-B PRECURSOR.
 (from "consensus-spt.pep")
 TOIG of: q9m4d4 check: 6480 from: 1 to: 148

ID Q9M4E4 PRELIMINARY; PRT; 148 AA.
 AC Q9M4E4;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE HORDOINDOLINE-B PRECURSOR.
 GN HORDOB.

OS Hordeum vulgare (barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_TaxId=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LEAF;
 RA Gauthier M.F., Cosson P., Guirao A., Alary R., Joudrier P.;
 RT "Puroindoline genes are highly conserved in diploid ancestor wheats
 RT and related species but absent in tetraploid Triticum species.";
 RL Plant Sci. 153:81-91(2000).
 DR EMBL: AJ24928; CAB89518.1; -
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereal_1tryp_1inh.
 DR Pfam: PF00234; tryp_alpha_1inh.
 DR PRINTS: PRO0808; AMLASEINHBR.
 DR SMART: SM00499; AAI; 1.
 KW Signal.
 FT CHAIN 1 19 POTENTIAL.
 FT SIGNAL 30 148 HORDOIINDOLINE-B
 SQ SEQUENCE 148 AA; 16656 MW; B4F8E3D098E9FAFC CRC64;
 Q9M4E4 Length: 148 September 12, 2002 12:10 Type: P Check: 6480
 Found using 'consensus' (consens.key)

1 MKTLFLIALIALVASTTFAQYSEVGVNNEVGAGSGSQCPLEPRKLSCDYMERCT
 40

61 MKFPYTWPTKMKKGGEHVEKCCQQLSQIAPQCRDSIRGMIGKLGFFGIRGVD

121 FKOIORAOSLPSKCMGADCKFSPSGYYW
 140

1 match found in sequence:
 q9sc07; PUROIINDOLINE B (FRAGMENT).
 (from "consensus-spt.pep")
 TOIG of: q9sc07 check: 6874 from: 1 to: 148
 ID Q9SC07; PRELIMINARY; PRT; 148 AA.
 AC Q9SC07;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE PUROIINDOLINE B (FRAGMENT).
 GN PINB-1.
 OS Triticum monococcum (Einkorn wheat) (small spelt).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Triticum.
 OC NCBI_TaxId=4568;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20092491; PubMed=10628869;
 RA Tranquilli G., Lijavetzky D., Muzzi G., Dubcovsky J.;
 RT "Genetic and physical characterization of grain texture-related loci
 RT in diploid wheat";
 RL Mol. Gen. Genet. 262:846-850(1999).
 DR EMBL: AJ242716; CAB63472.1; -
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereal_1tryp_1inh.
 DR Pfam: PF00234; tryp_alpha_1inh.
 DR PRINTS: PRO0808; AMLASEINHBR.
 DR SMART: SM00499; AAI; 1.
 FT NON_TER 1 1
 FT NON_TER 148 148

SQ SEQUENCE 148 AA; 16715 MW; B4E11DA9720CCEB9 CRC64;
 Q9SC07 Length: 148 September 12, 2002 12:10 Type: P Check: 6874
 Found using 'consensus' (consens.key)

1 MKTLFLIALIALVASTTFAQYSEVGVNNEVGAGSGSQCPLEPRKLSCDYMERCT
 40

61 MKFPYTWPTKMKKGGEHVEKCCQQLSQIAPQCRDSIRGMIGKLGFFGIRGVD

121 FKOIORAOSLPSKCMGADCKFSPSGYYW
 140

1 match found in sequence:
 q9scb8; BTL-CME2.1 PRECURSOR.
 (from "consensus-spt.pep")
 TOIG of: q9scb8 check: 8333 from: 1 to: 148
 ID Q9SCB8; PRELIMINARY; PRT; 148 AA.
 AC Q9SCB8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE BTL-CME2.1 PRECURSOR.
 GN ITRL.
 OS Hordeum vulgare (barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Triticeae; Hordeum.
 OC NCBI_TaxId=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. HATIF DE GRIGNON;
 RA ROYO J.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ251931; CAB64342.1; -
 DR HSSP: P01087; ITMO.
 DR InterPro: IPR003612; AAI.
 DR InterPro: IPR001768; Cereal_1tryp_1inh.
 DR Pfam: PF00234; tryp_alpha_1inh.
 DR PRINTS: PRO0808; AMLASEINHBR.
 DR SMART: SM00499; AAI; 1.
 KW Signal.
 FT CHAIN 1 24 POTENTIAL.
 FT SIGNAL 25 148 BTL-CME2.1.
 SQ SEQUENCE 148 AA; 16266 MW; 52A79ADC7E1F1573 CRC64;
 Q9SCB8 Length: 148 September 12, 2002 12:10 Type: P Check: 8333
 Found using 'consensus' (consens.key)

1 MAPKYOLISAAVMALIALVATATSFQDMCAPGDELPHNPLRACRTYVVSQSLCHQGRLLT
 29

61 SMDKRRCCDELAIPIYCRCEALRIITMGVVTWOGAFEGAYRDSFPCPREOTSTANL

121 VTPQECNLGTHGSATCPDELQPGYVL
 137

Times: -- Search Statistics --
 CPU 00:00:00.00
 Total Elapsed 00:00:01.00

Number of sequences searched:	33
Number of sequence hits:	33
Number of separate matches:	33
Number of sequence hits saved:	0

THIS PAGE BLANK (USPTO)

Kam 09/913,351

=> fil hcaplus

FILE 'HCAPLUS' ENTERED AT 11:29:06 ON 16 SEP 2002
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2002 AMERICAN CHEMICAL SOCIETY (ACS)

Copyright of the articles to which records in this database refer is held by the publishers listed in the PUBLISHER (PB) field (available for records published or updated in Chemical Abstracts after December 26, 1996), unless otherwise indicated in the original publications. The CA Lexicon is the copyrighted intellectual property of the the American Chemical Society and is provided to assist you in searching databases on STN. Any dissemination, distribution, copying, or storing of this information, without the prior written consent of CAS, is strictly prohibited.

FILE COVERS 1907 - 16 Sep 2002 VOL 137 ISS 12
FILE LAST UPDATED: 15 Sep 2002 (20020915/ED)

This file contains CAS Registry Numbers for easy and accurate substance identification.

CAS roles have been modified effective December 16, 2001. Please check your SDI profiles to see if they need to be revised. For information on CAS roles, enter HELP ROLES at an arrow prompt or use the CAS Roles thesaurus (/RL field) in this file.

=> d que 125

L1 251 SEA FILE=REGISTRY C[QERK].[MAGYFLTIWPV][LIVM].[15-39]CC[CSDENQR
KH][QEH][LV]/SQSP
L2 17562 SEA FILE=REGISTRY C.C..[LI]/SQSP
L3 126 SEA FILE=REGISTRY L1 AND L2
L4 76 SEA FILE=HCAPLUS L3
L5 442 SEA FILE=HCAPLUS DELGADO A?/AU
L6 5 SEA FILE=HCAPLUS VILLARRUBIA V?/AU
L7 23 SEA FILE=HCAPLUS GALLEG0 G?/AU
L8 2 SEA FILE=HCAPLUS TUDURI J?/AU
L9 0 SEA FILE=HCAPLUS L4 AND ((L5 OR L6 OR L7 OR L8))
L10 1 SEA FILE=HCAPLUS L4 AND GOMEZ?/AU
L11 0 SEA FILE=HCAPLUS L4 AND POLYSACCHARIDE#
L12 0 SEA FILE=HCAPLUS L4 AND MANNOS?
L13 0 SEA FILE=HCAPLUS L4 AND COMPLEX?
L14 6 SEA FILE=HCAPLUS L4 AND BIND?
L15 3 SEA FILE=HCAPLUS L4 AND CARBOHYDRATE?
L16 1 SEA FILE=HCAPLUS L14 AND PHOSPHATE?
L17 2 SEA FILE=HCAPLUS L4 AND GLYCO?
L18 4 SEA FILE=HCAPLUS L4 AND GLUCOS?
L19 0 SEA FILE=HCAPLUS L4 AND GALACTOS?
L20 1 SEA FILE=HCAPLUS L4 AND ASSOCIAT?
L21 16 SEA FILE=HCAPLUS (L9 OR L10 OR L11 OR L12 OR L13 OR L14 OR L15
OR L16 OR L17 OR L18 OR L19 OR L20)
L22 7 SEA FILE=HCAPLUS L4 AND ((MAJOR OR LARGE)(A)SUBUNIT#)
L23 9 SEA FILE=HCAPLUS L4 AND ((MINOR OR SMALL)(A)SUBUNIT#)
L24 5 SEA FILE=HCAPLUS L22 AND L23
L25 21 SEA FILE=HCAPLUS L21 OR L24

=> d ibib abs 125 1-21

THIS PAGE BLANK (USPTO)

L25 ANSWER 1 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2002:408695 HCAPLUS

DOCUMENT NUMBER: 136:385272

TITLE: A novel cocoa albumin and its use in the production of cocoa and chocolate flavour

INVENTOR(S): Kochhar, Sunil; Hansen, Carl Erik; Juillerat, Marcel
Alexandre; McCarthy, James

PATENT ASSIGNEE(S): Societe Des Produits Nestle S.A., Switz.

SOURCE: PCT Int. Appl., 25 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002042327	A2	20020530	WO 2001-EP13536	20011121
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG				

PRIORITY APPLN. INFO.: EP 2000-125523 A 20001121

AB A novel 2S cocoa albumin was isolated, purified and identified from cocoa beans. Enzymatic hydrolysis of the protein generated a pool of flavor precursors, peptides and amino acids that resulted in formation of cocoa flavor upon heating with sugars. The DNA encoding a precursor cocoa 2S protein was isolated from immature Theobroma cocoa seeds.

L25 ANSWER 2 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2002:332352 HCAPLUS

DOCUMENT NUMBER: 136:364897

TITLE: Expression cassettes using the myb11 promoter of Arabidopsis for tissue-specific expression of foreign genes in the embryonic epidermis and flower of plants
INVENTOR(S): Reindl, Andreas; Bischoff, Friedrich; Tonelli, Chiara; Petroni, Katia

PATENT ASSIGNEE(S): Basf Plant Science Gmbh, Germany

SOURCE: PCT Int. Appl., 64 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: German

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002034924	A2	20020502	WO 2001-EP12444	20011026
W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG,				

THIS PAGE BLANK (USPTO)

US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM
 RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY,
 DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF,
 BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG

DE 10053519 A1 20020502 DE 2000-10053519 20001027

PRIORITY APPLN. INFO.:

DE 2000-10053519 A 20001027

AB The invention relates to an expression cassette for expression of foreign genes in the embryonic epidermis and/or the flower of plants. The cassette uses the promoter of the myb11 gene of Arabidopsis thaliana or functional equiv. or equiv. fragments thereof that have substantially the same promoter activity, said promoter being operably linked with a nucleic acid sequence that is to be transgenically expressed. The invention further relates to vectors derived from said expression cassettes. The invention also relates to transgenic plants transformed with said expression cassettes or vectors, to cultures, parts or transgenic propagation material derived therefrom and to the use thereof for producing foodstuff, feedstuff, seeds, pharmaceuticals or fine chems.

L25 ANSWER 3 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2002:30491 HCAPLUS

DOCUMENT NUMBER: 136:397519

TITLE: Identification and molecular characterisation of hordoindolines from barley grain

AUTHOR(S): Darlington, H. F.; Rouster, J.; Hoffmann, L.; Halford, N. G.; Shewry, P. R.; Simpson, D. J.

CORPORATE SOURCE: IACR-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Bristol, BS41 9AF, UK

SOURCE: Plant Molecular Biology (2001), 47(6), 785-794

CODEN: PMBIDB; ISSN: 0167-4412

PUBLISHER: Kluwer Academic Publishers

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Grain texture in barley is an important quality character as soft-textured cultivars have better malting quality. In wheat, texture is considered to be detd. by the puroindolines, a group of basic hydrophobic proteins present on the surface of the starch granule. Hard wheats have been proposed to lack puroindoline a or to have mutant forms of puroindoline b which do not bind to the granule surface. Anal. of six barley cultivars (three soft-textured and three hard) showed that all contained proteins homologous to wheat puroindoline b, but PCR anal. failed to show any differences in amino acid sequences similar to those which have been proposed to det. textural differences in wheat. Southern blot anal. showed two hordoindoline b genes which were isolated and shown to encode proteins with 94% sequence identity. Expression of hordoindoline b mRNA occurred in the starchy endosperm and aleurone layer of the developing seed, but not in the embryo. Anal. of seven soft- and six hard-textured barley varieties showed that all contained hordoindoline a except two hard varieties (Sundance, Hart) which were subsequently shown to both lack hordoindoline a mRNA. It was therefore concluded that there is not a clear relationship between the presence of hordoindoline a and grain texture in barley.

REFERENCE COUNT: 34 THERE ARE 34 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L25 ANSWER 4 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2000:754713 HCAPLUS

DOCUMENT NUMBER: 133:330539

TITLE: Sequence-determined DNA fragments and corresponding

THIS PAGE BLANK (USPTO)

INVENTOR(S): encoded polypeptides from corn and Arabidopsis
 Alexandrov, Nickolai; Brover, Vyacheslav; Chen,
 Xianfeng; Subramanian, Gopalakrishnan; Troukhan, Maxim
 E.; Zheng, Liansheng; Dumas, J.

PATENT ASSIGNEE(S): Ceres Inc., USA

SOURCE: Eur. Pat. Appl., 339 pp.
 CODEN: EPXXDW

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 16

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 1033405	A2	20000906	EP 2000-301439	20000225
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				
CA 2300692	AA	20000825	CA 2000-2300692	20000225
CA 2302828	AA	20001006	CA 2000-2302828	20000406
EP 1055728	A2	20001129	EP 2000-303770	20000504
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				
EP 1054060	A2	20001122	EP 2000-304161	20000517
R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO				

PRIORITY APPLN. INFO.:

US 1999-121825P	P	19990225
US 1999-145918P	P	19990727
US 1999-145951P	P	19990728
US 1999-146386P	P	19990802
US 1999-146388P	P	19990802
US 1999-146389P	P	19990802
US 1999-147038P	P	19990803
US 1999-147204P	P	19990804
US 1999-147302P	P	19990804
US 1999-147192P	P	19990805
US 1999-147260P	P	19990805
US 1999-147303P	P	19990806
US 1999-147416P	P	19990806
US 1999-147493P	P	19990809
US 1999-147935P	P	19990809
US 1999-148171P	P	19990810
US 1999-148319P	P	19990811
US 1999-148341P	P	19990812
US 1999-148565P	P	19990813
US 1999-148684P	P	19990813
US 1999-123180P	P	19990305
US 1999-123548P	P	19990309
US 1999-125788P	P	19990323
US 1999-126264P	P	19990325
US 1999-126785P	P	19990329
US 1999-127462P	P	19990401
US 1999-128234P	P	19990406
US 1999-128714P	P	19990408
US 1999-129845P	P	19990416
US 1999-130077P	P	19990419
US 1999-130449P	P	19990421
US 1999-130510P	P	19990423
US 1999-130891P	P	19990423
US 1999-131449P	P	19990428

THIS PAGE BLANK (USPTO)

US 1999-132048P P 19990430
 US 1999-132407P P 19990430
 US 1999-132484P P 19990504
 US 1999-132485P P 19990505
 US 1999-132486P P 19990506
 US 1999-132487P P 19990506
 US 1999-132863P P 19990507
 US 1999-134256P P 19990511
 US 1999-134218P P 19990514
 US 1999-134219P P 19990514
 US 1999-134221P P 19990514
 US 1999-134370P P 19990514
 US 1999-134768P P 19990518
 US 1999-134941P P 19990519
 US 1999-135124P P 19990520
 US 1999-135353P P 19990521
 US 1999-135629P P 19990524
 US 1999-136021P P 19990525
 US 1999-136392P P 19990527
 US 1999-136782P P 19990528
 US 1999-137222P P 19990601
 US 1999-137528P P 19990603
 US 1999-137502P P 19990604
 US 1999-137724P P 19990607
 US 1999-138094P P 19990608

AB The present invention provides DNA mols. that constitute fragments of the genome and cDNAs from Zea mays mays (HYBRID SEED #35A19) and Arabidopsis thaliana (ecotype Wassilewski), and polypeptides encoded thereby. The DNA mols. are useful for specifying a gene product in cells, either as a promoter or as a protein coding sequence or as an UTR or as a 3' termination sequence, and are also useful in controlling the behavior of a gene in the chromosome, in controlling the expression of a gene or as tools for genetic mapping, recognizing or isolating identical or related DNA fragments, or identification of a particular individual organism, or for clustering of a group of organisms with a common trait. Arabidopsis DNA is used in the present expt., but the procedure is a general one. Protocols are provided for Southern hybridizations and transformation of carrot cells. [This abstr. record is one of 15 records supplemental to CA13316218528Q necessitated by the large no. of index entries required to fully index the document and publication system constraints.].

L25 ANSWER 5 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 2000:53704 HCAPLUS

DOCUMENT NUMBER: 132:103748

TITLE: Glutenin seed storage proteins modified to contain lipid or starch **binding** domains

INVENTOR(S): Appels, Rudi; Morell, Matthew; Bekes, Frank; Tamas, Laszlo

PATENT ASSIGNEE(S): Commonwealth Scientific and Industrial Research Organisation, Australia; Goodman Fielder Limited; Groupe Limagrain Pacific Pty. Ltd.

SOURCE: PCT Int. Appl., 77 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

THIS PAGE BLANK (USPTO)

```

-----
WO 2000002914      A1  20000120      WO 1999-AU563      19990712
W:  AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ,
    DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS,
    JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK,
    MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ,
    TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ,
    MD, RU, TJ, TM
RW:  GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW, AT, BE, CH, CY, DE, DK,
    ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG,
    CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG
AU 9947630          A1  20000201      AU 1999-47630      19990712
EP 1127066          A1  20010829      EP 1999-930937     19990712
R:   AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT,
    IE, SI, LT, LV, FI, RO
JP 2002520009      T2  20020709      JP 2000-559143     19990712
PRIORITY APPLN. INFO.:      AU 1998-4604      A  19980710
                                WO 1999-AU563      W  19990712
AB  A modified glutenin or seed-storage protein is modified by adding to the
    protein a domain which confers to the modified protein the ability to
    incorporate into gluten or bind a ligand or other macromol.
    Thus, the gene selected to construct vectors for such fusion proteins
    coded for C hordein, a storage protein from barley endosperm and
    characterized by an absence of cysteine residues. Lipid-binding
    domains were constructed from (1) barley oleosin, (2) wheat CM16 or CM17
    proteins, or (3) puroindoline A; a starch-binding domain was
    obtained from glucoamylase of Aspergillus niger. The modified fusion
    proteins were expressed in recombinant wheat. Such proteins have uses in
    the prepn. of food products or non-food products.
REFERENCE COUNT:      5      THERE ARE 5 CITED REFERENCES AVAILABLE FOR THIS
                                RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

```

L25 ANSWER 6 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1998:433562 HCAPLUS

DOCUMENT NUMBER: 129:215872

TITLE: Cloning and sequencing of a gene encoding a 2S albumin seed storage protein precursor from English walnut (*Juglans regia*), a major food allergen

AUTHOR(S): Teuber, Suzanne S.; Dandekar, Abhaya M.; Peterson, W. Richard; Sellers, Claudia L.

CORPORATE SOURCE: Department of Internal Medicine, University of California, Davis, CA, USA

SOURCE: Journal of Allergy and Clinical Immunology (1998), 101(6, Pt. 1), 807-814
CODEN: JACIBY; ISSN: 0091-6749

PUBLISHER: Mosby, Inc.

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Walnuts rank third in per capita consumption of tree nuts in the United States and can be assocd. with systemic IgE-mediated reactions in some individuals. The objectives of the study were to clone a gene encoding one of the major food allergens in the walnut kernel and to characterize the recombinant allergen. A cDNA expression library in the lambda vector Uni-ZAP, which was prepd. from walnut somatic embryos, was screened by using a patient's sera that reacted with multiple protein bands on immunoblotting. A cDNA clone contg. an insert of 663 bp was identified and named Jug r 1. DNA sequence anal. of this clone revealed that it encoded a protein 142 amino acids in length. Comparison of the encoded protein sequences with protein databases revealed that this clone exhibits

THIS PAGE BLANK (USPTO)

a 46.1% identity with the Brazil nut (*Bertholletia excelsa*) methionine-rich 2S albumin seed storage protein precursor, Ber e 1. Jug r 1 appears to be an important walnut food allergen; 12 of 16 sera from patients allergic to walnuts demonstrated IgE **binding** to the 2S albumin seed storage protein precursor fusion protein. An IgE-**binding** inhibition study suggests that the walnut 2S protein precursor undergoes posttranslational modification into a large and small subunit that is similar to castor seed, cottonseed, mustard seed, and Brazil nut 2S seed storage protein allergens. Interestingly, the gene encoding this allergenic protein in Brazil nuts has recently gained notoriety because of its exptl. use as a transgene to enhance the nutritional quality of legumes. This is now the sixth definitive 2S albumin seed storage protein demonstrated to **bind** IgE, suggesting that this class of proteins is inherently allergenic.

L25 ANSWER 7 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1997:506129 HCAPLUS
 DOCUMENT NUMBER: 127:175684
 TITLE: Amylase inhibitors
 INVENTOR(S): Miyazaki, Toshiyuki; Morimoto, Toshihisa; Murayama, Ryuji
 PATENT ASSIGNEE(S): Nisshin Flour Milling Co., Ltd., Japan; Nagata Sangyo Co., Ltd.
 SOURCE: Eur. Pat. Appl., 22 pp.
 CODEN: EPXXDW
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 785214	A2	19970723	EP 1997-100316	19970110
EP 785214	A3	19980909		
R: DE, FR, GB				
JP 09194392	A2	19970729	JP 1996-23446	19960118
AU 9710139	A1	19970724	AU 1997-10139	19970113
AU 712829	B2	19991118		
CA 2195051	AA	19970719	CA 1997-2195051	19970114
US 5726291	A	19980310	US 1997-782176	19970114

PRIORITY APPLN. INFO.: JP 1996-23446 19960118

AB Disclosed is an amylase inhibitor consisting of a protein constructed of 244 amino acid residues having two subunits each identified as SEQ ID NO:1, in which a single band is obsd. at a mobility of 0.26 by polyacrylamide gel electrophoresis, and having a high inhibitory activity against human pancreatic α -amylase to inhibit an increase in blood **glucose** level, control an insulin secretion or maintain the duration in a feeling of satiety for diet. The amylase inhibitor can be used alone or in combination with other known amylase inhibitors.

L25 ANSWER 8 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1997:503378 HCAPLUS
 DOCUMENT NUMBER: 127:161146
 TITLE: Amylase inhibitors of wheat for inhibiting accumulation of visceral fat
 INVENTOR(S): Miyazaki, Toshiyuki; Morimoto, Toshihisa; Murayama, Ryuji; Takase, Sachiko; Goda, Toshinao
 PATENT ASSIGNEE(S): Nisshin Flour Milling Co., Ltd., Japan; Nagata Sangyo Co., Ltd.

THIS PAGE BLANK (USPTO)

SOURCE: Eur. Pat. Appl., 22 pp.
 CODEN: EPXXDW
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 784978	A2	19970723	EP 1997-100399	19970111
EP 784978	A3	19980916		
R: DE, FR, GB				
JP 09194390	A2	19970729	JP 1996-23445	19960118
AU 9710117	A1	19970724	AU 1997-10117	19970110
AU 714225	B2	19991223		
CA 2195053	AA	19970719	CA 1997-2195053	19970114
US 5789380	A	19980804	US 1997-782177	19970114
PRIORITY APPLN. INFO.:			JP 1996-23445	19960118

AB An agent for inhibiting an accumulation of visceral fat is disclosed which comprises an amylase inhibitor of wheat origin as an active ingredient. Rats contg. 2.4% amylase inhibitor in their diet for a period of 2 wk consumed the same amt. of food and weighed the same at the end of the expt. as did the control group. However, the epididymal and mesenteric adipose tissue wt. was significantly less in the expt. group. Addnl., the activities the lipogenic enzymes fatty acid synthetase, malate dehydrogenase, and **glucose**-6-phosphate dehydrogenase were lower in the exptl. group.

L25 ANSWER 9 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1997:380351 HCAPLUS
 DOCUMENT NUMBER: 127:16885
 TITLE: Tea beverages containing amylase inhibitors for controlling obesity
 INVENTOR(S): Karita, Hanako; Yoshida, Mieko
 PATENT ASSIGNEE(S): Nisshin Flour Milling Co., Ltd., Japan
 SOURCE: Jpn. Kokai Tokkyo Koho, 10 pp.
 CODEN: JKXXAF
 DOCUMENT TYPE: Patent
 LANGUAGE: Japanese
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
JP 09094065	A2	19970408	JP 1995-273677	19950928

AB Amylase inhibitors such as proteins (amino acid sequence given) are added to wheat tea contg. little tannins, and the pH is adjusted to 3.8-5.4 with .gtoreq. 1 compd. selected from the group comprising lactic acid, citric acid, tartaric acid, malic acid, acetic acid, and ascorbic acid. The inhibitors are effective in controlling conversion of starch into **glucose**, and for treating diabetes, hyperlipidemia, and obesity.

L25 ANSWER 10 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1996:430176 HCAPLUS
 DOCUMENT NUMBER: 125:135744
 TITLE: Characterization of a sunflower seed albumin which **associates** with oil bodies
 AUTHOR(S): Thoyts, Patrick J. E.; Napier, Johnathan A.; Millichip, Mark; Stobart, A. Keith; Griffiths, W.

THIS PAGE BLANK (USPTO)

CORPORATE SOURCE: Trevor; Tatham, Arthur S.; Shewry, Peter R.
Department of Biochemistry, University of Bristol,
Bristol, BS8 1UG, UK
SOURCE: Plant Science (Shannon, Ireland) (1996), 118(2),
119-125
CODEN: PLSCE4; ISSN: 0168-9452
PUBLISHER: Elsevier
DOCUMENT TYPE: Journal
LANGUAGE: English
AB An Mr 22 500 protein was purified from isolated oil bodies from sunflower
cotyledons. N-terminal amino acid sequencing showed that this protein
belonged to the 2S albumin group of storage proteins. A corresponding
cDNA clone encoded a preproprotein, comprising two mature 2S albumin
proteins one of which corresponded to the oil-body assocd. albumin. In
contrast, the second albumin encoded by the cDNA was related to components
purified from a total albumin fraction by RP-HPLC.

L25 ANSWER 11 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1996:146882 HCAPLUS
DOCUMENT NUMBER: 124:278705
TITLE: Effect of chronic ingestion of .alpha.-amylase
inhibitor on the small intestinal mucosa of rats
AUTHOR(S): DiMagno, Eugene P.
CORPORATE SOURCE: GI Research Unit, Mayo Clinic, USA
SOURCE: Shoka to Kyushu (1995), 18(2), 51-5
CODEN: SHKYEZ; ISSN: 0389-3626
PUBLISHER: Nippon Shoka Kyushu Gakkai
DOCUMENT TYPE: Journal
LANGUAGE: Japanese
AB Effect of long-term inhibition of intraluminal .alpha.-amylase activity on
eating and the small intestinal mucosa was investigated in growing male
Wistar rats. The rats were fed the wheat amylase inhibitor (AI) for 21
days, and each AI rat was paired with control rats. Chronic ingestion of
the wheat AI significantly decreased food intake, reduced wt. gain and did
not alter stool output. After feeding of the wheat AI over 21 days, the
entire length of the small intestine was not changed. However, chronic AI
decreased the mucosal wt., protein and DNA contents of the upper small
intestine by reducing food intake and wt. gain (an overall nutritional
effect). By contrast, in the distal intestine, AI significantly increased
mucosal wt., RNA and activities of disaccharidases (sucrase, maltase and
lactase); this effect likely represents ileal adaptation to increased
amts. of undigested and unabsorbed **carbohydrate** delivered to the
distal intestine. These effect of the wheat AI suggest that it may be
useful to treat obesity, diabetes mellitus or both conditions.

L25 ANSWER 12 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1995:315692 HCAPLUS
DOCUMENT NUMBER: 122:100531
TITLE: Amylase inhibitors from wheat and their use for
controlling blood sugar level
INVENTOR(S): Miyazaki, Toshiyuki; Morimoto, Toshihisa; Murayama,
Ryuji; Matsubara, Hiroshi
PATENT ASSIGNEE(S): Nisshin Flour Milling Co., Ltd., Japan; Nagata Sangyo
Co., Ltd.
SOURCE: Eur. Pat. Appl., 21 pp.
CODEN: EPXXDW
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1

THIS PAGE BLANK (USPTO)

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
EP 618229	A1	19941005	EP 1994-302148	19940324
R: DE, FR, GB				
CA 2119660	AA	19940930	CA 1994-2119660	19940323
US 5444046	A	19950822	US 1994-216846	19940324
AU 9459072	A1	19941006	AU 1994-59072	19940325
AU 666431	B2	19960208		
JP 07041499	A2	19950210	JP 1994-79213	19940328
PRIORITY APPLN. INFO.:			JP 1993-91881	19930329
			JP 1993-148423	19930528

AB An amylase inhibitor consisting essentially of a protein constructed of 248 amino acid residues having two subunits, each identified as SEG ID NO:1, in which a single band is obsd. at a mobility of 0.26 by polyacrylamide gel electrophoresis. The amylase inhibitor is useful as an agent for inhibiting increase of blood **glucose** level, controlling insulin secretion or suppressing appetite, and as a food additive. For blood **glucose** level, controlling insulin secretion or suppressing appetite, and as a food additive. For the same use, the amylase inhibitor is used in combination with a protein composed of two subunits, each identified as SEQ ID NO:2, the total content of both proteins being not less than 20% by wt.

L25 ANSWER 13 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1994:319479 HCAPLUS

DOCUMENT NUMBER: 120:319479

TITLE: Vesicle transport and processing of the precursor to 2S albumin in pumpkin

AUTHOR(S): Hara-Hishimura, Ikuko; Takeuchi, Yuka; Inoue, Kaori; Nishimura, Mikio

CORPORATE SOURCE: Dep. Cell Biol., Natl. Inst. Basic Biol., Okazaki, 444, Japan

SOURCE: Plant Journal (1993), 4(5), 793-800

CODEN: PLJUED; ISSN: 0960-7412

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Cell fractionation of pulse-chase-labeled developing pumpkin cotyledons demonstrated that proprotein precursor to 2S albumin is transported from the endoplasmic reticulum to dense vesicles and then to the vacuoles, in which pro2S albumin is processed to the mature 2S albumin. Immunocytochem. anal. showed that dense vesicles of about 300 nm in diam. mediate the transport of pro2S albumin to the vacuoles. The primary structure of the precursor (16 578 Da) to pumpkin 2S albumin has been deduced from the nucleotide sequence of an isolated cDNA insert. The presence of a hydrophobic signal peptide at the N-terminus indicates that the precursor is a preproprotein that is converted into pro2S albumin after cleavage of the signal peptide. N-terminal sequencing of the pro2S albumin in the isolated vesicles revealed that the signal peptide is cleaved off co-translationally on the C-terminal side of alanine residue 22 of prepro2S albumin. By contrast, post-translational cleavages occur on the C-terminal sides of asparagine residues 35 and 74, which are conserved among precursors to 2S albumin from different plants. Hydrophathy anal. revealed that the two asparagine residues are located in the hydrophilic regions of pro2S albumin. These findings suggest that a vacuolar processing enzyme can recognize exposed asparagine residues on the mol. surface of pro2S albumin and cleave the peptide bond on the C-terminal side of each asparagine residue to produce mature 2S albumin in

THIS PAGE BLANK (USPTO)

the vacuoles.

L25 ANSWER 14 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1994:155222 HCAPLUS
 DOCUMENT NUMBER: 120:155222
 TITLE: Cloning of cDNA, expression and chromosomal location of genes encoding the three types of subunits of the barley tetrameric inhibitor of insect .alpha.-amylase
 AUTHOR(S): Medina, Joaquin; Hueros, Gregorio; Carbonero, Pilar
 CORPORATE SOURCE: E.T.S. Ing. Agron., UPM, Madrid, E-28040, Spain
 SOURCE: Plant Molecular Biology (1993), 23(3), 535-42
 CODEN: PMBIDB; ISSN: 0167-4412
 DOCUMENT TYPE: Journal
 LANGUAGE: English

AB Three cDNA clones from barley developing endosperm, corresponding to proteins BTAI-CMa, BTAI-CMb and BTAI-CMd, which are the three types of subunits of the tetrameric inhibitor of insect .alpha.-amylases, have been identified and sequenced. The deduced amino acid sequence of BTAI-CMb corresponds to the CM16/CM17 type of subunit in wheat (92/90% identical residues) and has on putative N-glycosylation site (NLT) and a possible kinase-C phosphorylation site (SCR). The BTAI-CMa sequence differs at four acid residues from a previously reported one from cv. Bomi and the sequence deduced for BTAI-CMd completes (11 N-terminal residues) and confirms previously available data. The gene for BTAI-CMa (Iat1) is located in the .beta. arm of barley chromosome 7H (syn.1), while genes for both BTAI-CMb (Iat2) and BTAI-CMd (Iat3) are in the long arm of chromosome 4H. The three genes are expressed in endosperm and their mRNAs are not detected in the other tissues tested, except Iat1, which seems to be expressed at a low level in coleoptile and roots, where it is switched off by 50 .mu.M Me jasmonate.

L25 ANSWER 15 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1992:1489 HCAPLUS
 DOCUMENT NUMBER: 116:1489
 TITLE: Isolation, characterization and expression of a gene coding for a 2S albumin from Bertholletia excelsa (Brazil nut)
 AUTHOR(S): Gander, E. S.; Holmstroem, K. O.; De Paiva, G. R.; De Castro, L. A. B.; Carneiro, M.; Grossi de Sa, M. F.
 CORPORATE SOURCE: Embrapa/Cenargen, Brasilia, 70.770, Brazil
 SOURCE: Plant Mol. Biol. (1991), 16(3), 437-48
 CODEN: PMBIDB; ISSN: 0167-4412
 DOCUMENT TYPE: Journal
 LANGUAGE: English

AB Two genes, BE2S1 and BE2S2, coding for methionine-rich albumins of Brazil nut (B. excelsa) have been cloned and their sequence detd. The genes are members of a multigene family and BE2S1 codes for one of the dominant 2S isoforms. Its expression is highly regulated during seed development and with respect to tissue specificity. Sequence anal. has shown that the genes contain one intron and that the promoter of BE2S1 shows a canonical TATA motif. The transcription initiation site is located 26 nucleotides downstream from the TATA box. Sequence comparison of the promoter regions of 2S genes from Brassica napus, Arabidopsis thaliana and B. excelsa revealed the presence of TCGA palindromic sequences that appear to be arranged in a 2S-specific manner.

L25 ANSWER 16 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1991:671965 HCAPLUS
 DOCUMENT NUMBER: 115:271965

THIS PAGE BLANK (USPTO)

TITLE: Characterization and evolution of napin-encoding genes in radish and related crucifers
 AUTHOR(S): Raynal, Monique; Depigny, Dominique; Grellet, Francoise; Delseny, Michel
 CORPORATE SOURCE: Lab. Physiol. Biol. Mol. Veg., Univ. Perpignan, Perpignan, 66860, Fr.
 SOURCE: Gene (1991), 99(1), 77-86
 CODEN: GENED6; ISSN: 0378-1119
 DOCUMENT TYPE: Journal
 LANGUAGE: English

AB Three cDNA clones, encoding napin storage proteins from radish, were isolated and sequenced. They fall into two classes differing in the size of the primary translation product. Sequences of the two classes are very well conserved and they display an organization very similar to that of the homologous genes from rapeseed and Arabidopsis which have previously been described. On the basis of hybridization intensity and the no. of restriction fragments, the authors est. that the radish napin multigene family is represented by eight to twelve members. The use of probes specific to each subfamily demonstrates that they contribute to a similar extent to the prodn. of napin mRNA. Anal. of the sequence data suggests that the napin ancestral genes are probably derived from successive duplication and divergence of a protogene. Comparing other available napin sequences with those of radish reveals intriguing features. Comparison of the coding sequences shows that the homol. between the radish and rapeseed sequences is much higher than that between each of the four members of the Arabidopsis gene family. This would suggest that the duplications which gave rise to the different members occurred independently in the two groups of species after sepn. of Arabidopsis from the Brassica lineage. However, a similar comparison carried out on the 3'-noncoding sequences does not support this hypothesis, but shows that slightly different duplicated genes probably already existed in the common ancestor to the three genera. This paradox can be resolved by assuming that, within each genus, coding sequences for napin-encoding genes have been considerably homogenized as a result of concerted evolution.

L25 ANSWER 17 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1990:566603 HCAPLUS
 DOCUMENT NUMBER: 113:166603
 TITLE: The Ricinus communis 2S albumin precursor: a single preproprotein may be processed into two different heterodimeric storage proteins
 AUTHOR(S): Irwin, Stephen D.; Keen, Jeffrey N.; Findlay, John B. C.; Lord, J. Michael
 CORPORATE SOURCE: Dep. Biol. Sci., Univ. Warwick, Coventry, CV4 7AL, UK
 SOURCE: MGG, Mol. Gen. Genet. (1990), 222(2-3), 400-8
 CODEN: MGGEAE; ISSN: 0026-8925
 DOCUMENT TYPE: Journal
 LANGUAGE: English

AB The R. communis (castor bean) 2S albumin is a heterodimer of glutamine-rich, disulfide-linked 4- and 7-kDa polypeptides. A cDNA library was constructed using mRNA from maturing castor bean endosperm as template. Clones contg. sequences complementary to albumin mRNA were isolated by hybridization using as a probe a mixt. of synthetic oligonucleotides representing sequences predicted for a peptide present in the 2S albumin **large subunit**. The nucleotide sequence contained an open reading frame encoding a preproprotein of 258 amino acid residues. The preproprotein included both polypeptides of the previously sequenced 2S albumin. In addn., this precursor included 2 further glutamine-rich sequences which, in terms of their size and conserved

THIS PAGE BLANK (USPTO)

cysteine residues typically found in seed proteins of the 2S albumin superfamily, possibly represent the small and large polypeptide subunits of a 2nd heterodimeric storage protein. A post-translational processing scheme is proposed which would result in a single preproprotein generating two distinct heterodimeric 2S albumins. The generation of a 2nd heterodimer seems likely since polypeptide candidates for its small and **large subunits** were found in the Ricinus 2S albumin fraction, and N-terminal protein sequencing confirmed the existence of the putative **small subunit**.

L25 ANSWER 18 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1990:456925 HCAPLUS

DOCUMENT NUMBER: 113:56925

TITLE: A method for the detection of IgE **binding** sequences of allergens based on a modification of epitope mapping

AUTHOR(S): Walsh, B. J.; Howden, M. E. H.

CORPORATE SOURCE: Sch. Chem., Macquarie Univ., 2109, Australia

SOURCE: J. Immunol. Methods (1989), 121(2), 275-80

CODEN: JIMMBG; ISSN: 0022-1759

DOCUMENT TYPE: Journal

LANGUAGE: English

AB An ELISA method for the rapid detn. of IgE-**binding** sites (allergenic determinants) of proteins is reported. The method utilizes the epitope mapping kit (H. Geysen et al., 1984) to synthesize hexapeptides of an allergen of interest, followed by a biotin-avidin system to detect peptide-bound IgE. The technique allows rapid localization of determinants from allergens of known sequence without the need to purify large amts. of allergen nor to generate peptides by cleavage of it. Using the results of the epitope mapping expts., a putative allergenic peptide contg. 18 amino acid residues from the sequence of a wheat allergen (.alpha.-amylase inhibitor) was identified and synthesized on polyamide resin. Testing of this peptide by radioallergosorbent test inhibition showed that it bound specific IgE in the sera of patients allergic to wheat.

L25 ANSWER 19 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1990:152810 HCAPLUS

DOCUMENT NUMBER: 112:152810

TITLE: The gene for trypsin inhibitor CMe is regulated in trans by the lys3a locus in the endosperm of barley (*Hordeum vulgare* L.)

AUTHOR(S): Rodriguez-Palenzuela, Pablo; Royo, Joaquin;

Gomez, Luis; Sanchez-Monge, Rosa; Salcedo, Gabriel; Molina-Cano, Jose Luis; Garcia-Olmedo, Francisco; Carbonero, Pilar

CORPORATE SOURCE: UPM, Madrid, E-28040, Spain

SOURCE: MGG, Mol. Gen. Genet. (1989), 219(3), 474-9

CODEN: MGGEAE; ISSN: 0026-8925

DOCUMENT TYPE: Journal

LANGUAGE: English

AB A cDNA encoding trypsin inhibitor CMe from barley endosperm was cloned and characterized. The longest open reading frame of the cloned cDNA codes for a typical signal peptide of 24 residues followed by a sequence which is identical to the known amino acid sequence of the inhibitor, except for an Ile/Leu substitution at position 59. Southern blot anal. of heat-barley addn. lines has shown that chromosome 3H of barley carries the gene for CMe. This protein is present at <2%-3% of the wild-type amt. in the mature endosperm of the mutant Ris.vphi. 1508 with respect to Bomi

THIS PAGE BLANK (USPTO)

barley, from which it has been derived, and the corresponding steady-state levels of the CMe mRNA are .apprx.1%. One or 2 copies of the CMe gene (synonym Itcl) per haploid genome have been estd. both in the wild type and in the mutant, and DNA restriction patterns are identical in both stocks, so neither a change in copy no. nor a major rearrangement of the structural gene account for the markedly decreased expression. The mutation at the lys3a locus in Ris.vphi. 1508 has been previously mapped in chromosome 7 (synonym 5H). A single dose of the wild-type allele at this locus (Lys 3a) restores the expression of gene CMe (allele CMe-1) in chromosome 3H to normal levels.

L25 ANSWER 20 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1989:129545 HCAPLUS

DOCUMENT NUMBER: 110:129545

TITLE: Determination of the processing sites of an Arabidopsis 2S albumin and characterization of the complete gene family

AUTHOR(S): Krebbers, Enno; Herdies, Lydia; De Clercq, Ann; Seurinck, Jef; Leemans, Jan; Van Damme, Jozef; Segura, Magdalena; Gheysen, Godelieve; Van Montagu, Marc; Vandekerckhove, Joel

CORPORATE SOURCE: Plant Genet. Syst., Ghent, B-9000, Belg.

SOURCE: Plant Physiol. (1988), 87(4), 859-66

CODEN: PLPHAY; ISSN: 0032-0889

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The most abundant isoform of the 2S albumin present in seeds of *A. thaliana* was sequenced and the corresponding gene isolated. Examn. of the protein and DNA sequences allows the detn. of the exact proteolytic cleavage sites during posttranslational processing. Like other 2S albumins, that of *Arabidopsis* is made as a prepropeptide. After removal of the signal peptide, the propeptide is cleaved at 4 other points, giving 2 subunits linked by a disulfide bridge(s). Comparison of these cleavage sites with those of 2S albumins of *Brassica napus* and *Bertholletia excelsa* suggests that while individual cleavages sites between species are conserved, the 4 processing sites within a species are not similar. This suggests that up to 4 different proteases may be involved in processing 2S albumins. The *Arabidopsis* 2S albumin gene was used to isolate the entire gene family. There are 4 genes, tightly linked in a tandem array. None of the genes contains an intron. Comparison of the predicted protein sequences shows that only 1 of the genes encodes the isoform detd. by protein anal. to be the most abundant, and therefore this gene is certain to be expressed. It is possible that some or all of the other 3 genes are also active.

L25 ANSWER 21 OF 21 HCAPLUS COPYRIGHT 2002 ACS

ACCESSION NUMBER: 1983:484230 HCAPLUS

DOCUMENT NUMBER: 99:84230

TITLE: The complete amino acid sequence of barley trypsin inhibitor

AUTHOR(S): Odani, Shoji; Koide, Takehiko; Ono, Teruo

CORPORATE SOURCE: Sch. Med., Niigata Univ., Niigata, 951, Japan

SOURCE: J. Biol. Chem. (1983), 258(13), 7998-8003

CODEN: JBCHA3; ISSN: 0021-9258

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The amino acid sequence of barley trypsin inhibitor was detd. The protein was a single polypeptide consisting of 121 amino acid residues and had a mol. wt. of 13,305. No free SH groups were detected by Ellman's reagent,

THIS PAGE BLANK (USPTO)

which indicated the presence of 5 disulfide bridges in the mol. The primary site of interaction with trypsin was tentatively assigned to the arginyl-leucyl residues at positions 33 and 34. On comparison of the sequence of this inhibitor with those of other proteinase inhibitors, it was found that barley trypsin inhibitor could not be classified into any of the established families of proteinase inhibitors and that this inhibitor should represent a new inhibitor family. On the other hand, this trypsin inhibitor showed considerable similarity to wheat .alpha.-amylase inhibitor throughout the whole sequence, suggesting a common ancestry for both proteins. This is the 1st case of a possible evolutionary relation between 2 inhibitors directed to totally different enzymes, a proteinase and a **glycosidase**.

THIS PAGE BLANK (USPTO)